

CHILD DEVELOPMENT

Editor

WILLIAM E. MARTIN
Purdue University

Consulting Editors

BETTYE McDONALD CALDWELL
Washington University School of Medicine

ERCEL S. EPPRIGHT
Iowa State College

EUGENE S. GOLLIN
Queens College

HARRY LEVIN
Cornell University

PAUL H. MUSSEN
University of California

KATHERINE E. ROBERTS
Merrill-Palmer School

WATARU W. SUTOW
University of Texas

WILLIAM W. WATTENBERG
Wayne State University

Managing Editor

KATE HOFFMAN



VOLUME 29

MARCH, 1958 to DECEMBER, 1958

CHILD DEVELOPMENT PUBLICATIONS

SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT, INC.

Purdue University, Lafayette, Indiana

1958

CONTENTS

VOLUME 29, 1958

BAER, DONALD M. (<i>see</i> Gewirtz, Jacob L.)	149
BALDWIN, ALFRED L., & LEVIN, HARRY. Effects of Public and Private Success or Failure on Children's Repetitive Motor Behavior	363
BALDWIN, ALFRED L. (<i>see also</i> Levin, Harry)	373
BELL, RICHARD Q. Retrospective Attitude Studies of Parent-Child Relations	323
BELL, RICHARD Q. (<i>see also</i> Schaefer, Earl S.)	339
BIJOU, SIDNEY W. A Child Study Laboratory on Wheels	425
BONNEY, MERL E., & NICHOLSON, ERTIE LOU. Comparative Social Adjustments of Elementary School Pupils with and without Preschool Training	125
BRACKBILL, YVONNE. Extinction of the Smiling Response in Infants as a Function of Reinforcement Schedule	115
BRACKBILL, YVONNE, & JACK, DONALD. Discrimination Learning in Children as a Function of Reinforcement Value	185
BURCHINAL, LEE G., GARDNER, BRUCE, & HAWKES, GLENN R. A Suggested Revision of Norms for the <i>Rogers Test of Personality Adjustment</i>	135
CHODACK, MARJORIE. (<i>see</i> Cohen, Walter)	473
COHEN, WALTER, HERSHKOWITZ, AARON, & CHODACK, MARJORIE. Size Judgment at Different Distances as a Function of Age Level	473
COHLER, J. ROBERT, JR. (<i>see</i> McNeil, Elton B.)	451
COLLIER, MARY J., & GAIER, EUGENE L. Adult Reactions to Preferred Childhood Stories	97
CRANDALL, VAUGHN, J., ORLEANS, SONYA, PRESTON, ANNE, & RABSON, ALICE. The Development of Social Compliance in Young Children	429
DAVIDSON, KENNETH S., SARASON, SEYMOUR B., LIGHTHALL, FREDERICK F., WAITE, RICHARD R., & SARNOFF, IRVING. Differences Between Mothers' and Fathers' Ratings of Low Anxious and High Anxious Children	155
DAVIDSON, KENNETH S. (<i>see also</i> Fox, Cynthia)	297
DAVIDSON, KENNETH S. (<i>see also</i> Sarason, Seymour B.)	105, 277, 287
DAVIS, JOHN M. A Reinterpretation of the Barker, Dembo, and Lewin Study of Frustration and Regression	503
DAVITZ, JOEL R. Contributions of Research with Children to a Theory of Maladjustment	3
EICHORN, DOROTHY H., & MCKEE, JOHN P. Physiological Instability During Adolescence	255
FITZSIMONS, RUTH. Developmental, Psychosocial, and Educational Factors in Children with Nonorganic Articulation Problems	481
FOX, CYNTHIA, DAVIDSON, KENNETH, LIGHTHALL, FREDERICK, WAITE, RICHARD, & SARASON, SEYMOUR B. Human Figure Drawings of High and Low Anxious Children	297

CHILD DEVELOPMENT

FRANKS, BEULAH B., & ROUSEY, CLYDE L. Visual Perception of Stutterers and Nonstutterers	445
GAIER, EUGENE L. (<i>see</i> Collier, Mary J.)	97
GARDNER, D. BRUCE, & SWIGER, MARYBELLE K. Developmental Status of Two Groups of Infants Released for Adoption	521
GARDNER, D. BRUCE. (<i>see also</i> Burchinal, Lee G.)	135
GARDNER, D. BRUCE. (<i>see also</i> Pease, Damaris)	141
GEWIRTZ, JACOB L., BAER, DONALD M., & ROTH, CHAYA H. A Note on the Similar Effects of Low Social Availability of an Adult and Brief Social Deprivation on Young Children's Behavior	149
HARTUP, WILLARD W. Nurturance and Nurturance-Withdrawal in Relation to the Dependency Behavior of Preschool Children	191
HAWKES, GLENN R. (<i>see</i> Burchinal, Lee G.)	135
HERSHKOWITZ, AARON. (<i>see</i> Cohen, Walter)	473
HORROCKS, JOHN E. (<i>see</i> Phelps, Harold R.)	69
HOUSE, BETTY J., & ZEAMAN, DAVID. A Comparison of Discrimination Learning in Normal and Mentally Defective Children	411
INSELBERG, RACHEL MARZAN. The Causation and Manifestations of Emotional Behavior in Filipino Children	249
JACK, DONALD. (<i>see</i> Brackbill, Yvonne)	185
JEFFREY, WENDELL E. Variables in Early Discrimination Learning: I. Motor Responses in the Training of a Left-Right Discrimination	269
JEFFREY, WENDELL E. Variables in Early Discrimination Learning: II. Mode of Response and Stimulus Difference in the Discrimination of Tonal Frequencies ..	531
JONES, MARY COVER, & MUSSEN, PAUL HENRY. Self-Conceptions, Motivations, and Interpersonal Attitudes of Early- and Late-Maturing Girls	491
JONES, MARY COVER. (<i>see also</i> Mussen, Paul Henry)	61
KAGAN, JEROME. Socialization of Aggression and the Perception of Parents in Fantasy	311
KAGAN, JEROME. (<i>see also</i> Mussen, Paul Henry)	57
KAUSLER, DONALD H. (<i>see</i> Trapp, E. Philip)	507
LETON, DONALD A. A Study of the Validity of Parent Attitude Measurement	515
LEVIN, HARRY, & BALDWIN, ALFRED L. The Choice to Exhibit	373
LEVIN, HARRY. (<i>see also</i> Baldwin, Alfred L.)	363
LEVY, NISSIM. A Short Form of the Children's Manifest Anxiety Scale	153
LIGHTHALL, FREDERICK F. (<i>see</i> Davidson, Kenneth S.)	155
LIGHTHALL, FREDERICK F. (<i>see also</i> Fox, Cynthia)	297
LIGHTHALL, FREDERICK F. (<i>see also</i> Sarason, Seymour B.)	105, 277, 287
LIPSITT, LEWIS P. A Self-Concept Scale for Children and Its Relationship to the Children's Form of the Manifest Anxiety Scale	463
McKEE, JOHN P. (<i>see</i> Eichorn, Dorothy H.)	255
McNEIL, ELTON B., & COHLER, J. ROBERT, JR. Adult Aggression in the Management of Disturbed Children	451
MARSHALL, HELEN R. Factors Relating to the Accuracy of Adult Leaders' Judgments of Social Acceptance in Community Youth Groups	417
MARSHALL, HELEN R. Prediction of Social Acceptance in Community Youth Groups ..	173
MEREDITH, HOWARD V. A Time Series Analysis of Growth in Nose Height During Childhood	19

INDEX

MILTON, G. A. A Factor Analytic Study of Child-Rearing Behaviors	381
MUSSEN, PAUL HENRY, & JONES, MARY COVER. The Behavior-Inferred Motivations of Late- and Early-Maturing Boys	61
MUSSEN, PAUL HENRY, & KAGAN, JEROME. Group Conformity and Perceptions of Parents	57
MUSSEN, PAUL HENRY. (see also Jones, Mary Cover)	491
NICHOLSON, ERTIE LOU. (see Bonney, Merl E.)	125
ORLEANS, SONYA. (see Crandall, Vaughn J.)	429
PEASE, DAMARIS, & GARDNER, D. BRUCE. Research on the Effects of Noncontinuous Mothering	141
PENNY, RONALD. Age and Sex Differences in Motivational Orientation to the Communicative Act	163
PERKINS, HUGH V. Factors Influencing Change in Children's Self-Concepts	221
PERKINS, HUGH V. Teachers' and Peers' Perceptions of Children's Self-Concepts	203
PHELPS, HAROLD R., & HORROCKS, JOHN E. Factors Influencing Informal Groups of Adolescents	69
PRESTON, ANNE. (see Crandall, Vaughn J.)	429
RABSON, ALICE. (see Crandall, Vaughn J.)	429
ROTH, CHAYA H. (see Gewirtz, Jacob L.)	149
ROUSEY, CLYDE L. (see Franks, Beulah B.)	445
SARASON, SEYMOUR B., DAVIDSON, KENNETH, LIGHTHALL, FREDERICK, & WAITE, RICHARD. Classroom Observations of High and Low Anxious Children	287
SARASON, SEYMOUR B., DAVIDSON, KENNETH, LIGHTHALL, FREDERICK, & WAITE, RICHARD. Rorschach Behavior and Performance of High and Low Anxious Children	277
SARASON, SEYMOUR B., DAVIDSON, KENNETH, LIGHTHALL, FREDERICK, & WAITE, RICHARD. A Test Anxiety Scale for Children	105
SARASON, SEYMOUR B. (see also Davidson, Kenneth S.)	155
SARASON, SEYMOUR B. (see also Fox, Cynthia)	297
SARNOFF, IRVING. (see Davidson, Kenneth S.)	155
SCHAEFER, EARL S., & BELL, RICHARD Q. Development of a Parental Attitude Research Instrument	339
SIEGEL, ALBERTA ENGVALL. The Influence of Violence in the Mass Media Upon Children's Role Expectations	35
SMOCK, CHARLES D. Perceptual Rigidity and Closure Phenomenon as a Function of Manifest Anxiety in Children	237
STEVENSON, HAROLD W., & STEWART, EDWARD C. A Developmental Study of Racial Awareness in Young Children	399
STEWART, BETTY RHEA. Developmental Differences in the Stability of Object Preferences and Conflict Behavior	9
STEWART, EDWARD C. (see Stevenson, Harold W.)	399
SWIGER, MARYBELLE K. (see Gardner, D. Bruce)	521
TEMPLIN, MILDRED C. General Information of Kindergarten Children: A Comparison with the Probst Study after 26 Years	87
TERRELL, GLENN. A Need for Simplicity in Research in Child Psychology	303
TERRELL, GLENN, JR. The Role of Incentive in Discrimination Learning in Children	231
TRAPP, E. PHILIP, & KAUSLER, DONALD H. Dominance Attitudes in Parents and Adult Avoidance Behavior in Young Children	507

CHILD DEVELOPMENT

WAITE, RICHARD R. (<i>see</i> Davidson, Kenneth S.)	155
WAITE, RICHARD R. (<i>see also</i> Fox, Cynthia)	297
WAITE, RICHARD R. (<i>see also</i> Sarason, Seymour B.)	105, 277, 287
WILSON, WILLIAM CODY. Imitation and the Learning of Incidental Cues by Preschool Children	393
ZEAMAN, DAVID. (<i>see</i> House, Betty J.)	411

CHILD DEVELOPMENT

Vol. 29
March, 1958
No. 1

Editor

WILLIAM E. MARTIN
Purdue University

Consulting Editors

BETTYE McDONALD CALDWELL
Washington University School of Medicine

ERCEL S. EPPRIGHT
Iowa State College

EUGENE S. GOLLIN
Queens College

HARRY LEVIN
Cornell University

PAUL H. MUSSEN
University of California

KATHERINE E. ROBERTS
Merrill-Palmer School

WATARU W. SUTOW
University of Texas

WILLIAM W. WATTENBERG
Wayne State University

Managing Editor

KATE HOFFMAN

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS
(of The Society for Research in Child Development, Inc.)
Purdue University
Lafayette, Indiana

Made in United States of America

Notice to Contributors

Child Development is devoted to basic research and theory concerning the growth and development of the child.

Manuscripts should be submitted in duplicate, each copy bearing only the title as a means of identification. The title, author's name, and institutional affiliation (together with any footnotes which reveal the identification of the author) should be submitted on separate pages.

Contributors are asked to refer to the *Publication Manual of the American Psychological Association* (1957 Revision) in the preparation of manuscripts.

The following rules require special emphasis:

1. All copy, including indented matter, should be typed double-spaced on white standard paper. Lines should not exceed six inches in length.

2. Footnotes should be few in number. They should be indicated in the text by superscript arabic numerals, numbered consecutively throughout the manuscript. The footnotes themselves should be typed double-spaced on a separate page placed at the end of the manuscript.

3. Tables and figures are costly to print and should be used only when essential. The author will be asked to contribute to the cost of composition of tabular and graphic material, which is excessive in amount or unduly complicated in form. Each table should be typed on a separate sheet. Insert a guide line (e.g., Table 1 about here) at the appropriate place in the manuscript. Figures should be prepared with black India ink, suitable for photography. Legends typewritten on an ordinary typewriter are unacceptable.

4. References in the text should be indicated by numbers in parentheses that refer to the list of references at the end of the article. Do not use footnotes for this purpose. In the list of references, the arrangement should be alphabetical by author's surname, or for multiple authors by the name of the senior; for any one author, by order of date. The following standard form should be observed:

Jones, H. E. Development in adolescence. New York: D. Appleton-Century, 1943.

Shepard, Winifred O., & Schaeffer, M. The effect of concept knowledge on discrimination learning. Child Developm., 1956, 27, 173-178.

The World List of abbreviations is used for journals.

The manuscript submitted should represent the final form in which the author wishes the paper to appear. All changes in proof (other than printer's errors) will be charged to the author.

All manuscripts should be addressed to William E. Martin, Editor, Child Development Publications, Purdue University, Lafayette, Indiana.

CONTRIBUTIONS OF RESEARCH WITH CHILDREN TO A THEORY OF MALADJUSTMENT¹

JOEL R. DAVITZ

Teachers College, Columbia University

The major theories of psychological maladjustment, such as those of Freud and Horney, agree that childhood experience is of primary importance in the development of neurosis. It is remarkable, therefore, that almost all of these theories were originally formulated without reference to any but the most casual kind of research about children. In the main, the meanings of childhood experiences were interpreted by these theorists largely on the basis of inferences from observations of adult behavior. Only *after* the theories had been formulated did researchers try to find support for these chronologically "backward looking" inferences in observations of children. Unfortunately, these efforts have not been consistently successful. In fact, very little evidence supports most of the central hypotheses about childhood and maladjustment proposed by the major theories of neurosis. Consider, for example, the striking absence of reliable research evidence in support of the Freudian hypotheses about the Oedipus complex or about the effects of infantile frustrations on personality development. The absence of such support some 50 years after these hypotheses were published suggests that another approach to a theory of maladjustment might possibly result in a more substantial relationship between theory and research evidence. If we accept the premise that childhood experience is important in the development of maladjustment, rather than begin with observations of adults and infer back to childhood, perhaps it might be more useful to begin building a theory of maladjustment by starting with observations of children who are maladjusted.

The relationship between a child's early experiences with key figures in his environment and his subsequent personality development has been the focus of an impressive amount of research in child psychology. There are indeed large gaps of knowledge in this area, but the research available at the present time does permit some rather broad generalizations relevant

¹ Paper read at symposium on Contributions of Research with Children to Behavior Theory, Amer. Psychol. Ass., Chicago, September, 1956.

to a theory of maladjustment. Consider, for example, Symonds' studies (12) of parent-child relationships. On the basis of extensive research and a thorough review of the evidence, he concluded that if one or both parents strongly reject a child, he tends to become fearful, hostile, overaggressive, and may exhibit antisocial behaviors such as lying and stealing. In another series of researches, Redl and Wineman (8) reported that children who experienced severe rejection and abuse exhibit fear and anxiety, low frustration tolerance, irresponsibility, a tendency to panic when confronted with a novel situation, and greater difficulty in reacting to failure. Research also indicates that these negative reactions acquired in the home environment generalize to other interpersonal situations. Skeels and his co-workers (11) at Iowa, for example, in their study of environmental stimulation concluded that children reared in an orphanage without warm and affectionate parent-child relations tend to acquire attitudes of distrust and defensiveness which generalize in their relationships with others outside of the orphanage. And studies by Van Houten (13), Johnson (6), Hattwick and Stowell (4), and others have demonstrated that fearful responses acquired in the parent-child relationship tend to generalize to the school situation. From these and other related studies which have obtained similar results, we may conclude that in a punishing, rejecting interpersonal environment, the child tends to become fearful, and may exhibit hostility, overaggressiveness, distrust, irresponsibility, and other asocial or antisocial behaviors. Furthermore, the evidence suggests that these reactions tend to generalize and to influence the child's behavior in a wide range of interpersonal situations.

Up to this point we have been concerned with a straightforward stimulus-response course of learning. Punish, reject, abuse a child, and he will become fearful, suspicious, hostile. We are now faced with somewhat more complex problems. Although our research findings overwhelmingly support the common-sense generalization that threat engenders fear, not all children react fearfully to apparently equivalent punishing experiences. Some seem to blissfully live through years of threat and become happy, healthy persons. Others become maladjusted with seemingly minimal punishing experience. These contradictory observations present a crucial problem, and we do not have a satisfactory solution. But at least part of the solution may be found in studies of individual differences among infants in emotional sensitivity and responsiveness. The dimensions of temperamental differences in infancy, which are, presumably, genetically determined, have not yet been fully defined or measured; however, even at the present time there is a substantial body of research which supports the notion that such differences in sensitivity and responsiveness could well account for children's differential reactions to interpersonal threat. Shirley (10), for example, has described a wide range of differences in timorousness and irritability in early infancy that persist later in the child's life. Certainly it does not seem unreasonable to predict that the child who is

physically more sensitive, whose emotional reactions are cued off more easily and with greater intensity, will respond differently to stress in his interpersonal environment than will a child who is relatively insensitive to emotional stimulation.

A second problem involves the persistence of the child's emotional response long after the initial learning experience. Why doesn't the fear extinguish? This problem is what Mowrer (7) has termed the "neurotic paradox." A child whose parents are consistently rejecting and punishing learns to fear his parents and generalizes this fear to other parental figures. Sooner or later, however, one would expect this fear to decrease and in the normal course of events to extinguish. In the life of a psychologically healthy child, probably many fears are learned and then gradually extinguish. But in the maladjusted child, and adult, the fear of others persists for long periods of time without any apparent reduction or extinction. In fact, persistent, long-term fear might well be the chief defining characteristic of the maladjusted person.

The solution of this problem stems from the interpersonal nature of the child's learning experiences. The child lives and behaves in relationship to other people, and his behavior elicits some response from these other people. The infant's cry, for example, elicits some response from another person and this response may be either rewarding or punishing to the child. Sears (9) has developed this view of child development in his description of early socialization, emphasizing the interaction effects of parents and children behaving in relationship to each other.

Now consider the kinds of defensive behaviors found to be characteristic of children who have experienced severe punishment and rejection. Redl and Wineman (8) found hate, irresponsibility; Symonds (12) found hostility, overaggressiveness, truancy, lying, stealing; other common characteristics that have been found are suspicion, severe withdrawal, various forms of delinquency, extreme attention-getting, and a long list of similar kinds of behaviors. What kinds of responses would one expect a child's persistent hostility, overaggressiveness, extreme attention-getting, severe withdrawal to elicit from other persons in his environment? In most cases, we could safely predict that the child's behavior will elicit some form of counterpunishment or rejection. Studies by Anderson (1), by Bishop (2), Brewer (3), and others support this prediction. Briefly, these researches have demonstrated that integrative behavior in one person tends to induce integrative behavior in others; domination tends to incite domination; aggression tends to be met with aggression. Anderson (1), in his review of the research with children in regard to this phenomenon, has called this "circular behavior." Jackson and Todd (5) in their discussion of maladjusted children have called this the "vicious circle."

The answer, therefore, to the question of why the maladjusted child's fear doesn't extinguish is simply that the fear continues to be reinforced. The sequence of events resulting in continued reinforcement of the child's

fear begins when the child learns to fear others as a consequence of rejection and punishment. Having learned to fear other persons, he develops defensive reactions that temporarily reduce his fear; but, in the long run, these defensive reactions elicit responses from other persons which are punishing and which reinforce his fear of others. For example, if the child resorts to hostile defenses, he tends to elicit hostility in others. If he becomes severely withdrawn, he tends to elicit withdrawing or rejecting reactions in others. The kinds of negative, defensive behaviors which typically characterize the maladjusted child are precisely the kinds of behaviors which tend to elicit punishing reactions from others in the child's environment. Thus, other persons continue to punish the maladjusted child, and his fear of others persists. Punishment and rejection give rise to fear; fear promotes defensive reactions; and the defensive reactions elicit further punishment. This is the vicious circle of maladjustment. The child lives in an interpersonal world of constant threat. The consequence is unhappiness and fear.

Having started with observations of parent-child interactions, and following, at each step, the implications of research primarily with children, we have arrived at the beginnings, at least, of a theory of children's maladjustment without positing unconscious motives, id and superego conflicts, basic hostilities, unconscious guilts, and the host of intervening variables and hypothetical constructs prevalent in current clinical usage. Nevertheless, the sequence of events we have outlined accounts for the central feature of psychological maladjustment, viz., long-term, persistent fear. This sequence of events can be considered only the beginnings of a theory or an approach to the problem of maladjustment. There are many questions to be investigated before we may profitably build a more detailed and complex theoretical structure. For example, what are the dimensions of emotional sensitivity in the infant, and what are the characteristics of the interaction between the genetically determined sensitivities of the child and the stresses of his interpersonal environment? What are the factors which determine whether or not a child will resort to defenses which elicit threat and punishment from others? Do maladjusted children consistently perceive key figures in their environment as threatening? And do maladjusted children consistently exhibit the kinds of behaviors which tend to elicit threat, rejection, and punishment from others? If the sequence of events we have outlined is a valid description of children's maladjustment, can it also account for maladjustment in adults? Do neurotic adults, for example, perceive their interpersonal environments as threatening, and do they behave in just those ways that tend to reinforce this threat? If the central cause of a person's maladjustment is to be found in his interpersonal relations, what are the psychological implications of various ethical systems and codes of interpersonal behavior?

Our examination of research with children in this attempt to formulate the beginnings of a theory of psychological maladjustment has led to an

interpretation of childhood experience that differs from those theories which started with observations of adults and worked their ways back to childhood by imaginative inference. Notwithstanding the provocative value of inferences about childhood from adult behavior, increased understanding of maladjustment in children depends primarily upon further research focused on the behavior and experience of the child.

REFERENCES

1. ANDERSON, H. H., & ANDERSON, G. L. Social development. In L. Carmichael (Ed.), *Manual of child psychology* (2nd Ed.). New York: Wiley, 1954. Ch. 19.
2. BISHOP, B. M. Mother-child interaction and the social behavior of children. *Psychol. Monogr.*, 1951, 65 (Whole No. 328).
3. BREWER, J. E. The measurement of the behavior of second-grade children in relation to the teachers' dominative and socially integrative contacts. Unpublished doctoral dissertation, Univer. of Illinois, 1940.
4. HATTWICK, B. W., & SOWELL, M. Relation of parental over-attentiveness to children's work habits and social adjustment. *J. educ. Res.*, 1936, 30, 169-176.
5. JACKSON, L., & TODD, K. M. *Child treatment and the therapy of play*. (2nd Ed.). London: Methuen, 1948.
6. JOHNSON, A. M., et al. School phobia. *Amer. J. Orthopsychiat.*, 1937, 11, 702-711.
7. MOWRER, O. H. Learning theory and the neurotic paradox. *Amer. J. Orthopsychiat.*, 1948, 18, 571-610.
8. REDL, F., & WINEMAN, D. *Children who hate*. Glencoe: Free Press, 1951.
9. SEARS, R. R. A theoretical framework for personality and social behavior. *Amer. Psychologist*, 1951, 6, 476-483.
10. SHIRLEY, M. M. *The first two years*: Vol. III. *Personality manifestations*. Minneapolis: Univer. of Minnesota Press, 1933.
11. SKEELS, H. M., et al. A study of environmental stimulation: An orphanage preschool project. *Univer. Iowa Stud. Child Welfare*, 1936, 1, 156.
12. SYMONDS, P. M. *The dynamics of parent-child relationships*. New York: Bureau of Publications, Teachers Coll., Columbia Univer., 1949.
13. VAN HOUTEN, J. Mother-child relationships in twelve cases of school phobia. *Smith Coll. Stud. Soc. Work*, 1947-48, 18, 161-180.

DEVELOPMENTAL DIFFERENCES IN THE STABILITY OF OBJECT PREFERENCES AND CONFLICT BEHAVIOR¹

BETTY RHEA STEWART

University of Kansas²

While children have served as subjects in a number of conflict and choice-making studies, the focus has not been on their performance, but on the testing of general conflict theory. However, the conceptual picture of the young child presented by Baldwin (1), Lewin (5, 6, 7, 8), and Werner (11) suggests that differences in psychological maturity, roughly defined as increasing differentiation in the life space, could be expected to affect the child's stability of object preference and his behavior in a conflict situation.

The young child, as they depict him, lives in a here-and-now world where perception, emotion, action, and general cognitive activities are still partially fused. As the child grows older, differentiation between and within these capacities develops. He has an extended range of emotions instead of his earlier diffuse reactions. His perceptions are less affectively tinged and are well articulated rather than global. He has a past and a future which are significant influences in his present. He is able to withhold action until he has grasped a situation cognitively without being forced to act by the pull of the immediate situation.

The study was designed to test the effects of increasing age upon stability of object preferences and upon behavior in a conflict situation of the approach-approach type. The following hypotheses, based on the theoretical framework offered by Baldwin, Lewin, and Werner, were made about differences in the behavior of the young preschool age child and the older child.

1. The young child's preferences for objects will be less stable than the older child's, both in an immediate situation and after an interval of time.

As Werner portrays the young child, his behavior is governed largely by his affective needs. These needs, according to Lewin (5), shift often, causing alterations in the attractiveness of objects in the child's environ-

¹ This study was performed as part of a dissertation submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy in the Department of Psychology, University of Kansas. The writer is indebted to Dr. Edward L. Wike for his direction of the research and to Dr. Anthony J. Smith and Dr. Roger Barker for their assistance.

² Now at Department of Child Study, Vassar College.

ment. Therefore, in expressing preference for one object rather than another, the young child may change his choice within a comparatively short time without realizing the inconsistency of his action. Even greater disparity may be predicted if the time interval between choices is long, since his time perspective is limited.

The older child, in contrast, should be more consistent in his choices. According to Lewinian theory, the psychological world of the child becomes more stable and more differentiated with increasing age, so the older child should be less at the mercy of his immediate needs and the environmental forces of the moment. In addition, his time perspective is increasing, thus making it possible for his past actions to have implications for his present behavior.

2. Young children will agree less among themselves on their preferences for objects than will older children.

Lewin lists as one of the principal differences in developmental stages "an increasing *differentiation* of every level of the life space into a multitude of social relations and areas of activities" (7, p. 797). The young child does not have a clearly defined age and sex role. He is still at liberty to express a large range of preferences which do not have to conform to those of other children in his group. The older child's preferences, however, are more often imposed on him by the expectations of adults and his peers. A boy, for example, will incur ridicule if he displays interest in dolls and a girl will be criticized if she is too active in boys' games.

3. In an approach-approach conflict situation the young child will make a decision more quickly, but less decisively, than the older child.

Werner describes the young child's world as one of action, where perception and direct motor activity often function as a single unit. The young child, seeing two desirable objects side by side, can be expected, according to this view, to react to them with some kind of direct motor behavior. But both Werner and Lewin indicate that while the young child may act impulsively in a choice situation, he will not necessarily act decisively.

The older child, with his greater self-control in the face of internal and external pressures, should take longer to make his choices, and should make them with more certainty.

4. In an approach-approach conflict situation the young child's choices will be less influenced by differing valences of objects, as ranked by himself, than will the older child's choices.

The young child may experience the presentation of each set of choice objects as a total situation and not as related to previous choices from a larger group of objects. The young child's world has not yet expanded logically, according to Baldwin, so it is difficult for him to perceive how isolated aspects of a situation fit together to form a total pattern.

The older child, with his longer time perspective and his greater ability to comprehend interrelationships, may be influenced in his present choices by his earlier statements of preference for the objects.

METHOD

Subjects

The younger group was composed of 30 children from the University of Kansas nursery school, ranging from 3 years to 5 years, 4 months, with an average age of 4 years. There were 16 girls and 14 boys in the sample. As a group they were intellectually superior with a mean IQ of 131 on the Stanford-Binet test, Form M, and a range from 108 to 152.

The older group was made up of 30 fourth graders from a Lawrence, Kansas, public school. Their average age was 9 years, 10 months, with a range from 9 years, 3 months, to 10 years, 4 months. There were 14 girls and 16 boys in the group. Their mean IQ, based on a combination of the California Mental Maturity Tests for elementary grades, and the Stanford-Binet, Form L, was 124, with a low score of 103 and a high score of 151.

Materials

Materials used in the experiment were 12 ten cent toys: doll, comb and mirror set, jacks, marbles, ray gun, tank, pencil, paper pad, paints, telescope, puzzles, and toy snake.

Procedure

Ranking procedure. Pilot study results indicated that young children cannot rank toys in the usual way. The present method gave a system of preferential ranking of the toys which the younger group could understand and perform.

Each child was shown a 20 by 15 inch box with a hinged lid, containing all the toys arranged in a standard order. Although the child was not allowed to become involved in a game with the toys, his attention was called to all of them, and he was encouraged to explore their possibilities. When he had examined the toys, he was told to look at *all* the toys carefully and to pick out the one toy he liked best. The child's choice was timed with a stop watch from the moment the instructions were completed until he made his choice verbally, by pointing, or by picking up the toy. The toy was then removed from view and the lid of the box was closed until the child was again told to choose the toy he liked best of the ones remaining. The procedure was repeated and times were recorded until all the toys were chosen.

Paired comparisons procedure. Twelve combinations of toys were arranged from the three toys the child had ranked as his top choices and the three toys that he had liked least. The child was presented either with a pair of toys he had ranked high or a pair he had ranked low; he was never given a pair made up of one high and one low ranking toy. The order of presentation of the paired comparisons was randomized so that no subjects in a sample received the toys in exactly the same order. Each toy appeared once on the right and once on the left. In addition, the order

of appearance for most and least liked pairs of toys was randomized for each subject.

Immediately after the ranking procedure a 20 by 20 inch cardboard screen was placed on the table between *S* and *E*. The child was told that two toys would be placed behind the screen. When *E* lifted the screen, *S* was to look at both toys carefully and to indicate the toy he liked "the very best." The screen was placed about six inches from the edge of the table and the toys were put just behind the screen approximately 12 inches apart. Choices were timed with a stop watch and choice time was defined as the period of time extending from the instant the screen was lifted until a final choice was indicated. But times were also noted if *S* grabbed both toys or pointed quickly from one toy to the other before making his final choice.

Ranking procedure retest. From 8 to 11 days later subjects in both groups were again asked to rank the toys, the same procedure being used as in the first ranking situation. This completed the experiment with the younger children. The older children were interviewed briefly about their general reactions to the toys.

RESULTS

Mean scores, ordinarily the appropriate measures of central tendency for such data as the individual ranking times and paired comparisons choice times, would have been misleading, since the raw data for many of the subjects were markedly skewed. Median ranking times and choice times for each subject were employed throughout the study as the best representatives of individual scores. Rank order correlations between the first and second ranking orders, computed for each subject, were treated as indices of stability of preference and were treated as individual scores in subsequent analyses. Because the frequency distributions of median choice times and of the rank order correlations were also skewed, the results throughout were analyzed by nonparametric statistics.

Both groups were significantly consistent in choices from first to second rankings, as indicated by a mean rank order correlation of .41 for the nursery school group and .80 for the grade school group. However, in the frequency distributions of rhos for the two groups, the young children's scores fell into a roughly bimodal distribution, with some children evidencing little stability of choice from one ranking to the other, while others were highly consistent. The grade school children's scores were concentrated at the upper end of the distribution. Wilcoxon's *T* test for unpaired replicates, applied to rhos of the two groups, yielded a critical ratio of 4.91, significant at far less than the .001 level. This finding strongly supported the hypothesis that the older children would show greater stability of object preferences over an extended period of time than the younger children.

To test the stability of choice behavior for young and older children when only a short period of time elapsed, comparisons were made between preferences expressed on the first ranking and those indicated immediately afterward in the paired comparisons procedure. The data were categorized as highly consistent, consistent, and inconsistent. If, in the paired comparisons situation, the child chose in all or all but two of his 12 choices the toy he had ranked higher a few minutes before, he was considered highly consistent. If he chose the toy with the higher ranking between half and three-fourths of the time, he was categorized as consistent. His behavior was scored inconsistent if he chose a toy he had ranked lower more than half of the time. A chi square of 14.72 (2 *df*), significant at less than the .001 level, indicated that nursery school children were much less consistent in their choices than grade school children when a short period of time separated expressions of preference.

It was hypothesized that members of the younger group would differ widely from one another in their rankings of the 12 toys, while the preferential rankings of the older children were expected to be quite consistent from individual to individual. To test this assumption, Kendall's coefficient of concordance (*W*) was computed for each group's first order of ranking and on the data for each group broken down according to sex. The significance of the *W*'s was tested by transforming them into *F*s and evaluating the *F* distribution according to a method described by Walker and Lev (10). Results of the analysis are presented in Table 1.

It is evident from the table that the fourth grade group agreed only a little better among themselves on their preferences than the younger children. But, when the data were analyzed according to sex for the two groups, the seeming lack of agreement among the older group appeared to be an artifact caused by sex differences in toy preferences. Since the grade school boys were displaying a consistent order of preference for certain toys while the girls were showing a consistent order for different toys, the real stability of the preferences was concealed by the statistical

TABLE I
COEFFICIENTS OF CONCORDANCE FOR TOY PREFERENCES

Group	<i>W</i>	<i>df</i>	<i>F</i>
Nursery School Total	.205	330	7.48*
Grade School Total	.233	330	8.79*
Nursery School Boys	.301	154	5.60*
Nursery School Girls	.215	176	4.11*
Grade School Boys	.599	176	22.41*
Grade School Girls	.465	154	11.30*

* Significant at .001 level.

CHILD DEVELOPMENT

method. There were also indications that the boys in both groups agreed more in expressions of preference than the girls, though this cannot be stated conclusively, since there is no significant test of differences between *W*'s available.

A comparison of ranking times for the nursery school group and the fourth graders disclosed that the younger children chose more rapidly than the older children when confronted with the whole array of toys in the first ranking situation. Differences according to Wilcoxon's *T* were significant at the .01 level, the nursery school group having a median of 2.42 seconds and the grade school children a median of 3.16 seconds. Differences between the groups were not significant on second ranking times.

When the total choice times for the two groups were compared for the paired comparisons situation, differences between the groups were significant at the .001 level, the nursery school group taking much longer to make their choices than the grade school children. Median time for the nursery school choices was 3.76 seconds, while the grade school children had a median time of 2.07 seconds. This finding was in direct contrast to the original hypothesis that the young children would react faster than the older ones in a choice situation.

TABLE 2
COMPARISON OF CHOICE TIMES FOR NURSERY SCHOOL CHILDREN AND
GRADE SCHOOL CHILDREN ON MOST LIKED TOYS AND
LEAST LIKED TOYS

<i>Group</i>	<i>Valence of Toys</i>	<i>Median Choice Time</i>	<i>Difference</i>
WITHIN GROUPS			
Nursery School	Most Liked	3.76	
	Least Liked	3.64	.12
Grade School	Most Liked	1.94	
	Least Liked	2.09	.15*
BETWEEN GROUPS			
Nursery School	Most Liked	3.76	
Grade School	Most Liked	1.94	1.82†
BETWEEN GROUPS			
Nursery School	Least Liked	3.64	
Grade School	Least Liked	2.09	1.55†

* Significant at .002 level.

† Significant at .001 level.

The second part of the hypothesis stated that young children would be more indecisive than older ones. Choice behavior was considered indecisive (a) if the child grabbed both toys and had to be reminded that he could choose only one, (b) if he verbally expressed a desire to have both toys and did not choose one, or (c) if he wavered between the two toys, first indicating one and then the other. Seven of the 15 nursery school children displaying indecisive behavior grabbed both toys, while none of the four grade school children categorized as indecisive did this. All indecisive behavior of the grade school children was shown, instead, by a wavering between the two choice objects, with no verbally expressed desires for both. Four of the nursery school children, in contrast, said they wanted both, and four others indicated one toy, then the other in their choices. A chi square of 11.14 (1 *df*), significant at the .001 level, supported the hypothesis that the younger children would be more indecisive.

As Table 2 indicates, the younger subjects chose toys in the paired comparisons situation as rapidly when they had ranked them high as when they had ranked them low, a sign test computed on the difference between median choice times for the most and least liked toys indicating nonsignificance. The older children, on the other hand, took significantly longer to choose toys they had ranked low than those they had ranked high, a sign test indicating differences significant at the .002 level.

When comparisons were made between the two groups rather than within each group the difference between median choice times on the most liked toys was significant at the .001 level as tested by Wilcoxon's *T*, with the nursery school children taking longer. The nursery school children were also slower than the older group in choices between toys they liked least, the difference again being significant at the .001 level.

DISCUSSION

The design of the experiment corresponded closely to an approach-approach conflict situation, since, theoretically, choices were to be made between objects of varying degrees of positive valence or between objects about which the subjects did not feel strongly. None of the choices was intended to cause active avoidance-avoidance reactions, as is the case, for example, in animal experiments where shock is utilized. However, the behavior of some of the older children indicated that valences of least liked toys were negative enough at times to represent a mild avoidance-avoidance conflict.

The choice situation for the younger children more nearly resembled an approach-approach situation for all the toys. However, if the approach-approach quality of the conflict situation were the only factor operating, the young children's choice times for *all* the toys should have been approximately the same as for the older group's choice times on the *most liked* toys, an approach-approach situation also. According to conflict theory,

the longest choice times for either group should have been the older group's time on the least liked toys, an avoidance-avoidance conflict. But the nursery school children were significantly slower in all their choices, which would indicate that the level of psychological maturity must be taken into account also in making predictions about children's conflict behavior.

The hypothesis that young children would react faster than older children to the choice materials was the only major prediction not supported by the data. Studies have shown that young children's reaction times *are slower than older children and adults* (3), but observations of the nursery school group's choice making behavior indicated that the process of deciding was in itself difficult for them. Subjects who grabbed both toys often had to be told more than once that they could choose only one toy and even then they had difficulty reaching a final decision. Contrary to Lewin's theory of conflict, which states that the region of indecision is less attractive when the conflict is between objects of positive valence, the young children seemed to find it more satisfying to remain in an indecisive state than to give up definitely the possibility of having both toys. Barker (2) suggests that in an approach-approach conflict, where the valences are both highly positive and approximately equal, a relatively greater increase in the strength of one force will be necessary to resolve the conflict than where the two objects are intrinsically less attractive, but still positive. The intensity of the young children's conflicts in choosing between two liked toys may have been stronger than the intensity of the older children's conflicts to the extent that the toys were intrinsically more attractive for the younger than the older children.

That young children in general may have a more undifferentiated reaction to valence was suggested by their inability to rank in the usual way. The concept of an order of preference, where over-all feeling of liking or disliking an object are secondary to the relationship between the attractiveness of one toy and another, appeared to be a very difficult idea for the young child to grasp. In several instances the nursery school children had difficulty making selections even in the modified ranking situation because they could not conceive of liking any single toy best in a group where they disliked all the toys from which the choice was to be made.

SUMMARY AND CONCLUSIONS

The experiment was designed to examine the effects of increasing age upon stability of object preferences and behavior in conflict situations of an approach-approach type, chronological age being the independent variable. Psychological maturity and valence were assumed to be intervening variables. The dependent variables were ranking times, choice times between paired comparisons of the toys, and stability of preferences over periods of a few minutes and approximately a week.

The two samples consisted of 30 nursery school children and 30 fourth grade school children. In the first part of the experiment subjects were shown 12 ten cent toys which they ranked by a modified ranking procedure. In the second phase the children chose between pairs of most liked and least liked toys as determined by their previous rankings. The third part consisted of a retest of ranking behavior after an 8 to 11 day lapse, and a brief interview with the older children. The following conclusions were drawn from the results:

1. Young children are less consistent in their choices than older children both in an immediate choice situation and after an interval of time.
2. Young children differ more from each other in their preferences than do grade school children.
3. Young children choose more rapidly than older children when given a whole array of toys to choose from, but less rapidly when the task is to choose one of two toys.
4. The behavior of young children in a choice situation is more indecisive than older children's behavior.
5. The choice behavior of young children is not influenced by the valence of the objects, as ranked by themselves, but older children find it more difficult to make choices between objects they dislike.

REFERENCES

1. BALDWIN, A. L. *Behavior and development in childhood*. New York: Dryden, 1955.
2. BARKER, R. G. An experimental study of conflict by children; time elapsing and amount of vicarious trial and error behavior occurring. In Q. McNemar & M. A. Merrill (Eds.), *Studies of personality*. New York: McGraw-Hill, 1942.
3. JONES, H. E. Reaction time and motor development. *Amer. J. Psychol.*, 1937, 50, 181-194.
4. KENDALL, M. G. *Rank correlation methods*. London: Charles Griffin, 1948.
5. LEWIN, K. Environmental forces in child behavior and development. In Carl Murchison (Ed.), *A handbook of child psychology*. Worcester, Mass.: Clark Univer. Press, 1931.
6. LEWIN, K. The conceptual representation and measurement of psychological forces. In Donald K. Adams & Helge Lundholm (Eds.), *Contributions to psychological theory*. Durham, N.C.: Duke Univer. Press, 1938.
7. LEWIN, K. Behavior and development as a function of the total situation. In Leonard Carmichael (Ed.), *Manual of child psychology*. Worcester, Mass.: Clark Univer. Press, 1946.

CHILD DEVELOPMENT

8. LEWIN, K. Regression, retrogression, and development. In Dorwin Cartwright (Ed.), *Field theory in social science*. New York: Harper, 1951.
9. MILLER, N. E. Experimental studies in conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*, Vol. I. New York: Ronald Press, 1944.
10. WALKER, HELEN M., & LEV, J. *Statistical inference*. New York: Henry Holt, 1953.
11. WERNER, H. *Comparative psychology of mental development*. New York: Follett, 1948.
12. WILCOXON, F. *Some rapid approximate statistical procedures*. New York: American Cyanamid Company, 1949.

A TIME SERIES ANALYSIS OF GROWTH IN NOSE HEIGHT DURING CHILDHOOD¹

HOWARD V. MEREDITH

Iowa Child Welfare Research Station, State University of Iowa

This study pertains to the growth of the individual child. It is related historically to the pioneering longitudinal studies of Montbeillard (23) in the eighteenth century, Bowditch (4) in the nineteenth century, and Baldwin (3) in the early years of the present century. At closer range, it is a part of a program of research on individual growth fostered at the State University of Iowa during the past two decades (16).

In the pursuit of biologically valid knowledge on individual growth it is important to select for study measures that can be obtained with very high dependability (16); skeletal nose height is such a measure. Serial data for this dimension, spaced at yearly intervals, have been amassed and analyzed.

The findings are organized under three headings: (1) trend, (2) velocity, and (3) relation with other variables. Primary emphasis is placed on the classification of magnitude trends, the analysis of velocity differences, and the comparison of body traits with respect to growth rate correlations and increment trends.²

SUBJECTS

The subjects were 80 children, 38 boys and 42 girls, residing in or near Iowa City, Iowa. All were voluntary participants in the Facial Growth Study, a long-term research program at the State University of Iowa begun in 1946 under the joint sponsorship of the Department of Orthodontics and the Iowa Child Welfare Research Station. Enrollment in the study was based on likelihood of continuing residence in the community and

¹ Supported in part by a research grant, D-217, from the National Institute for Dental Research, of the National Institutes of Health, U.S. Public Health Service.

² Using terminology appropriate to the view that growth is "a form of motion," Tanner (26) refers to magnitude or integral trends as "distance" curves, and to increment or differential trends as "velocity" curves.

CHILD DEVELOPMENT

willingness to participate; it was not related to odontic condition or orthodontic need.

The sample comprised a group of American-born white children predominantly of northwest European ancestry and above average in socioeconomic status. Every subject was examined annually from age 5 years to age 12 years. Forty-five of the subjects, 19 boys and 26 girls, were examined also at ages 13 years and 14 years. The examinations were made on or near each child's birthday anniversary.

DATA

Data for skeletal nose height were derived from *norma lateralis* roentgenograms of the face. The roentgenograms were obtained with (a) the head maintained in a Higley (9) positioner, (b) the film paralleling the sagittal plane near the side of the head, (c) the distance from the X-ray source to the sagittal plane held constant at 150 cm., and (d) the central ray passing through the acoustic meatuses.

To minimize age-to-age variation in landmark determination, a given landmark was registered at one sitting on the entire series of roentgenograms for a subject. For example, two persons experienced in roentgenographic anthropometry examined the nasion region of the films for Subject F1 and reached agreement on the location of nasion. This was done with the aid of a magnifying glass, and the point decided upon for each film was pricked with a needle probe. Employing the same procedure, the most anterior point of the anterior nasal spine was marked.

The roentgenographic distance from the superior landmark (nasion) to the inferior landmark (tip of anterior nasal spine) was measured by means of a steel millimeter tape read to the nearest 0.01 cm. Films were taken for measurement in random order, and each was measured by two anthropometrists working independently. When the two independent records were found to agree within 0.02 cm., they were averaged. In instances of greater discrepancy (there were less than 10 per cent of these) two additional measurements were taken and the mean of the four determined.

These measures derived from the films were corrected for roentgenographic enlargement and blurring. A previous publication (21) describes in detail the methods employed to obtain the corrections. It will suffice here to record that the roentgenographic measures were converted to measures of skeletal nose height by subtraction of amounts varying from 0.23 cm. to 0.41 cm.

Summarily, meticulous care was taken to obtain valid measures of skeletal nose height. Attention was given to both the reliability of the data (by controls to minimize chance error in subject orientation, landmark location, and measurement) and their relevance (by corrections for the systematic errors that roentgenography entailed).

FINDINGS FOR TREND

The Group Trend

Table 1 presents means for skeletal nose height derived from the measures on those subjects studied annually from age 5 years to age 12 years. It will be seen that the means for boys and girls (a) do not differ at any age by more than 0.06 cm. and (b) increase with age in a highly similar manner. In view of these findings, attention will be turned to the central tendency trend for the two sexes combined.

TABLE 1

ARITHMETIC MEANS FOR SKELETAL NOSE HEIGHT (NASION TO ANTERIOR NASAL SPINE IN CENTIMETERS) ON 80 NORTH AMERICAN WHITE CHILDREN STUDIED LONGITUDINALLY

Age (yrs.)	Boys (N = 38)	Girls (N = 42)	Both Sexes (N = 80)
5	3.95*	3.89	3.92
6	4.11	4.06	4.08
7	4.25	4.20	4.22
8	4.37	4.33	4.35
9	4.49	4.46	4.48
10	4.60	4.57	4.59
11	4.70	4.68	4.69
12	4.79	4.77	4.78

* The reader may wish to interpret the sample means for a particular age as parameter estimates. Standard errors of the mean are 0.03 cm. for boys 5 years of age, 0.02 cm. for girls 5 years of age, and 0.03 cm. to 0.04 cm. in all other instances.

Figure 1 shows that the curve drawn to the composite means for both sexes from Table 1 has an upward slope of diminishing steepness, i.e., the group magnitude trend for skeletal nose height over the childhood period between ages 5 years and 12 years is a negatively accelerated increasing function of age. An equation for this trend, obtained by the method of least squares, is: Skeletal Nose Height = $2.710 (\text{Age})^{0.2285}$. Goodness of fit is satisfactory; the difference between comparable means, estimated and obtained, in no instance exceeds 0.01 cm.

Central tendency trends of the same general form (increase during childhood at a declining rate) have been obtained for bizygomatic width of the upper face (14) and bigonial width of the lower face (21). The former study extends over the childhood period 4 to 10 years and the latter over the period 5 to 11 years.

To this point, the group trend prior to age 12 years has been considered. Although the size of the sample drops from 80 to 45 after age 12 years,



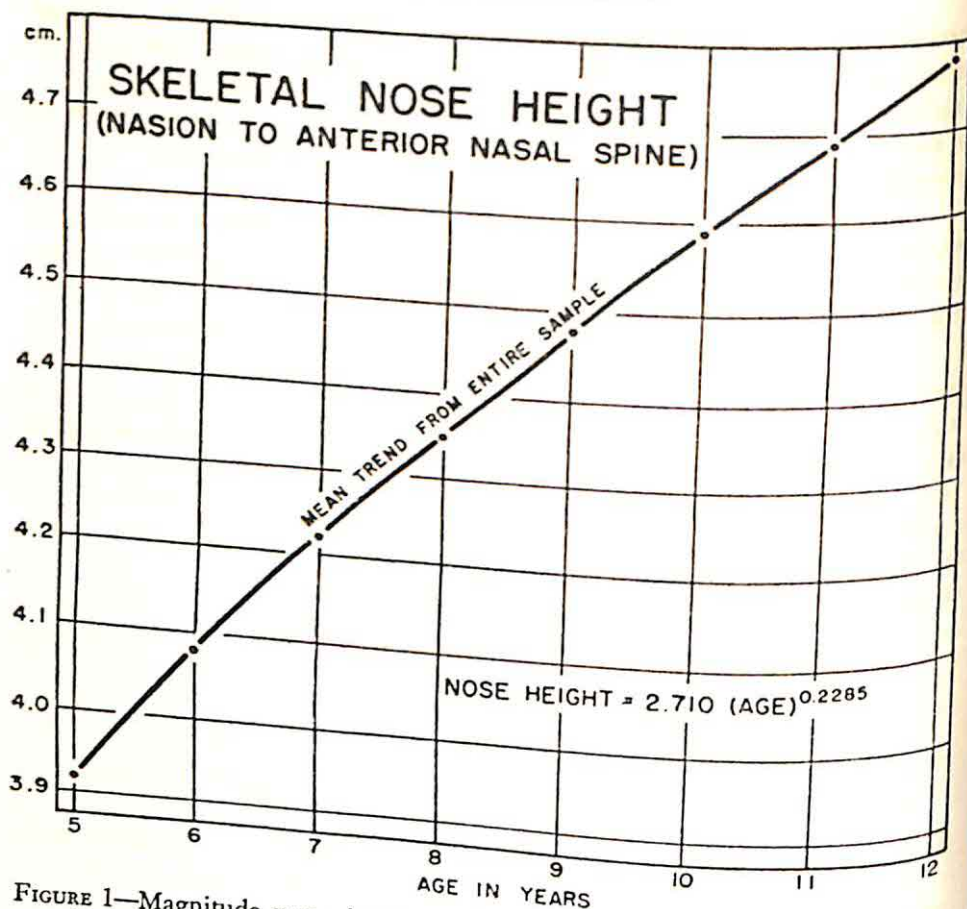


FIGURE 1—Magnitude curve drawn to group means for 80 North American white children of both sexes studied longitudinally.

some inquiry up to age 14 years is possible. Means on this smaller sample were obtained at ages 12, 13, and 14 years. The mean at age 12 years is 4.78 cm., the same as that for the full sample of 80 children (*see* Table 1). At the two succeeding ages the means are 4.88 cm. and 4.97 cm., respectively. These figures, together with the values from Table 1 for ages 10, 11, and 12 years, warrant the inference that between ages 10 and 14 years the group trend for skeletal nose height approximates a slowly rising rectilinear function of age. A similar group trend for this age span has been found from numerous investigations on width of the upper face (14).

The means of Table 1, in common with those reported in a few other investigations on both living and non-living children (1, 12), are for skeletal nose height. Another group of studies on the living child reports means based on direct measurement of the nose from the level of nasion to subnasale, i.e., means for integumental nose height (5, 6, 7, 8). To convert the values from one approach to equivalent values for the other approach necessitates information on the comparability of results when the two

methods are employed on the same group of subjects. Measures of integumental nose height at age 5 years, obtained using the Ashley-Montagu (2) procedure for locating the level of nasion, were available for the 80 subjects of the present study. The mean of these measures (4.13 cm.) exceeds its corresponding value for skeletal nose height by 0.21 cm.

Returning to findings specific for sex, although the Table 1 means on boys are slightly higher than those on girls the sex difference is not statistically significant. Nor can it be posited on the basis of previous studies that there is a sex difference in nose height during the childhood years (1, 5, 7, 25, 28). However, the availability of additional subjects at age 5 years made it possible to test for sex differences in skeletal nose height on samples of 82 boys and 84 girls. These enlarged samples gave means of 3.98 cm. and 3.86 cm. for boys and girls, respectively. The difference is significant at a high level of confidence ($t = 3.58$, df 164), and harmonizes with a substantial body of knowledge for other facial dimensions showing that on the average boys are larger than girls. Variability findings from these enlarged samples at age 5 years are as follows: boys, standard deviation 0.23 cm. and coefficient of variation 5.7 per cent; girls, standard deviation 0.20 cm. and coefficient of variation 5.2 per cent. The obtained differences in variability are nonsignificant statistically.

Trends for the Individual

The procedure used in analyzing individual magnitude trends for skeletal nose height was as follows:

1. The seriatim data from each subject was plotted on a Cartesian grid. As is customary, ages were used as abscissa values and measures of skeletal nose height as ordinate values.
 2. A curve was drawn connecting each series of plotted points, yielding a total of 80 curves. Individually, these curves showed clear expression of trend, i.e., there was little manifestation of random or cyclic fluctuation (16, 21, 24). Utilizing a phrase from Quenouille (22), the time series on each child consisted "largely of a trend."
 3. The curves were classified with respect to similarity of trend. That is, rising rectilinear trends were grouped in one class, trends that ascended in a path concave to the age base line in another class, and so forth.
 4. A composite trend representing each grouping of individual curves was derived. This was done by tabulating the measures on each member of a class according to age, computing means for the class at successive yearly ages, plotting these class means, and drawing a curve to them.
- Findings will be presented first for the larger sample studied between ages 5 and 12 years, then for the smaller sample studied between ages 5 and 14 years.

For the period from age 5 years to age 12 years, three types of individual magnitude trend were found for skeletal nose height. The commonest type is an ascending curvilinear trend concave downward; this

CHILD DEVELOPMENT

represents more than two-thirds of the 80 subjects studied. The other types are: an ascending rectilinear trend, and an ascending curvilinear trend concave downward in the first half of the period and convex downward in the last half. Figure 2 portrays the composite trends derived to delineate the three types, and indicates the absolute and relative frequencies of each type in the sample studied.

The lower right panel of Figure 2 displays two individual trends illustrating the range of variation within the concave class. Although both

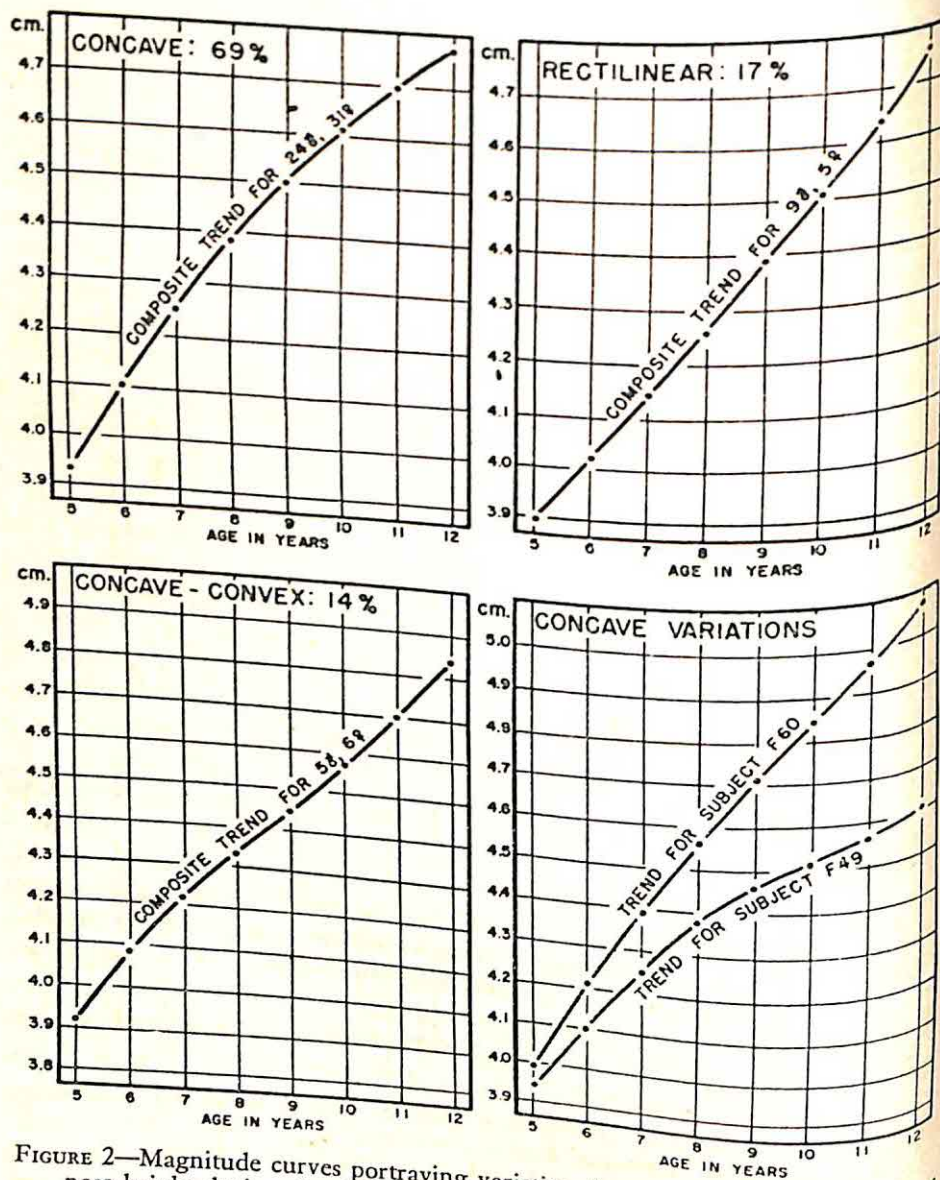


FIGURE 2—Magnitude curves portraying varieties of individual growth in skeletal nose height during the childhood period between 5 and 12 years of age.

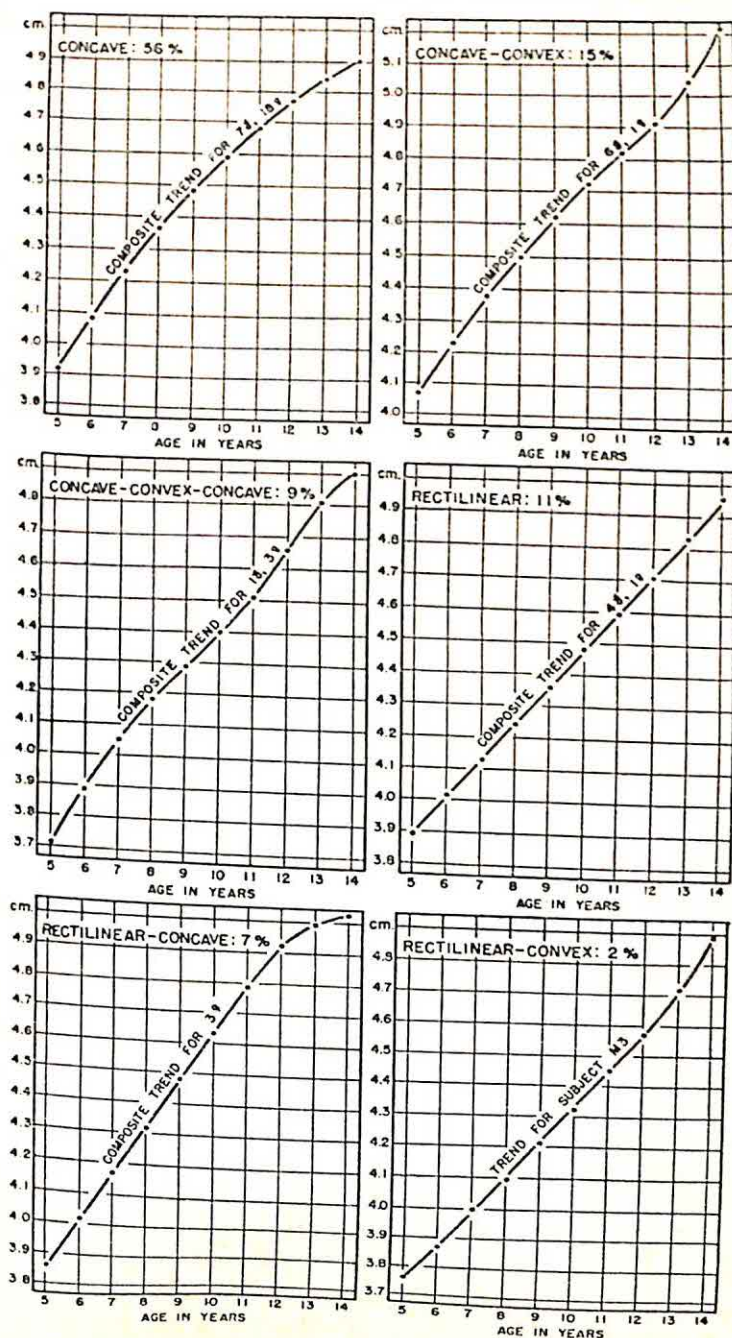


FIGURE 3—Magnitude curves depicting types of growth in skeletal nose height between 5 and 14 years of age.

trends manifest an upward course of declining steepness, the over-all rise is 0.41 cm. more for Subject F60 than for F49. Further consideration of individual differences in velocity will comprise a later section of the paper.

Figure 3 depicts each of the six categories of magnitude trend found for the period between age 5 years and age 14 years. Examination of this graph shows:

1. Twenty-five of the 45 individual curves for skeletal nose height ascend at a progressively declining rate. In other words, more than one-half of the children studied to age 14 years exhibit a rising trend concave to the age base line.

2. Eleven of the individual curves for skeletal nose height are ascending trends of diminishing steepness during the first part of the age span only. For seven, this early period of downward concavity is followed by a span of downward convexity; for four, it is followed by first downward convexity and then downward concavity. Projecting into the period between age 14 years and early adulthood, it appears reasonable to assume that the "concave-convex" trends are incomplete "concave-convex-concave" trends.

3. The remaining individual curves for skeletal nose height are rectilinear during at least the first one-half of the period. Five are rising straight-line trends out to 14 years of age, three are negatively accelerated trends between ages 12 and 14 years, and one is a positively accelerated trend after age 12 years. Extrapolating again (16), it may reasonably be inferred that (a) to the "rectilinear" trends there will be added either a segment that is concave downward, or a segment that is first convex downward and then concave downward, and (b) to the "rectilinear-convex" trend there will be added a negatively accelerated (concave) segment.

Synthesizing, magnitude trends representing individual growth in skeletal nose height during childhood and early adolescence have been graphed (Figures 2 and 3) and discussed. These trends in part support and in part suggest the generalized inference: Between 5 years of age and early adulthood, white children of northwest European ancestry travel along one of four trend paths in augmenting their nose height. Explicitly,

1. For about 1 child in 2 the magnitude trend of nose height is a negatively accelerated increasing function of age.

2. For about 1 child in 4 the magnitude trend of nose height is a two-cycle increasing function of age, i.e., there is a period of positive velocity between two periods of negative velocity.

3. For about 1 child in 8 the magnitude trend of nose height is a rising rectilinear function of age during the childhood years and a negatively accelerated function of age during adolescence.

4. For an occasional child the magnitude trend of nose height proceeds at a constant absolute rate during childhood, at an increasing absolute rate in early adolescence, and at a declining absolute rate in late adolescence.

It should be recognized that the foregoing findings stem from the Cartesian grid approach to the study of individual growth. From the semilog-

arithmic grid approach, which also is useful in biometric research, somewhat different trend groupings would be expected (21).

FINDINGS FOR VELOCITY

Individual and Age Differences

Variability of velocity in skeletal nose height was investigated on the 80 subjects studied from age 5 years to age 12 years. Centimeter and percentage increments were determined for four annual intervals and the full seven-year period. Each series of derived values was arranged in order of magnitude. Statistical reductions for these 10 ordered series are assembled in Table 2. Selected findings are:

1. There is a general decrease in velocity of growth in nose height with age. In centimeter terms, the minimum increase during the sixth postnatal year coincides with the 30th percentile of the increment distribution for the twelfth year; the 70th percentile of the distribution for the fifth year coincides with the maximum increase during the twelfth year; and so forth. In percentage terms, the typical gain for the sixth year (4.0 per cent) is twice that for the twelfth year, and the terminal values of the annual percentage gain distributions decline from 2.1-6.1 to 0.7-3.9.
2. Over the seven-year interval between age 5 years and age 12 years, increase in nose height for some children is double that for other children.

TABLE 2

VARIABILITY OF INCREASE IN SKELETAL NOSE HEIGHT FOR FOUR ANNUAL PERIODS AND A SEVEN-YEAR PERIOD ON 80 NORTH AMERICAN WHITE CHILDREN OF BOTH SEXES

Age Interval (yrs.)	P E R C E N T I L E S :					
	Minimum	10	30	70	90	Maximum
C E N T I M E T E R I N C R E A S E						
5 - 6	0.08	0.12	0.14	0.17	0.20	0.24
7 - 8	0.06	0.10	0.11	0.14	0.16	0.19
9 - 10	0.05	0.08	0.10	0.12	0.14	0.17
11 - 12	0.03	0.05	0.08	0.11	0.14	0.17
5 - 12	0.55	0.71	0.79	0.91	1.01	1.15
P E R C E N T A G E I N C R E A S E *						
5 - 6	2.1	3.0	3.6	4.4	5.0	6.1
7 - 8	1.3	2.3	2.7	3.3	3.8	4.5
9 - 10	1.1	1.7	2.2	2.8	3.2	4.0
11 - 12	0.7	1.1	1.7	2.4	2.9	3.9
5 - 12	14.5	17.7	19.8	23.8	25.9	30.2

* Computed as: Absolute increase in skeletal nose height during the age interval indicated $\times 100$ / Magnitude of skeletal nose height at the age the interval begins.

The obtained absolute and relative dispersions extend from 0.55 cm. to 1.15 cm., and from 14.5 per cent to 30.2 per cent. For the absolute increments as a series, 1 in 10 lies below 0.71 cm., 1 in 4 between 0.79 cm. and 0.91 cm., and 1 in 10 above 1.01 cm. Corresponding subdivision values for the percentage increment series are 17.7, 19.8, 23.8 and 25.9, respectively.

Correlation Findings

Two varieties of relationship involving rate of growth in skeletal nose height were studied; they may be designated (a) size-gain and (b) gain-gain. The Pearson product-moment method r was used; all analyses were based on data for 80 subjects. Size-gain findings are:

1. The correlation is close to zero ($r = -0.04$) for the association between nose height at age 5 years and centimeter increase in nose height over the age span 5 to 12 years.

2. There is low negative correlation ($r = -0.35$, df 78, $t = 3.72$) denoting the association between nose height at age 5 years and percentage increase in nose height over the age span 5 to 12 years.

Several other studies have reported size-gain correlation coefficients for facial dimensions. These studies encompass: bizygomatic width of the upper face, size at 4 years with gain from 4 to 9 years (14); bigonial width of the lower face, size at 5 years with gain from 5 to 11 years (21); dental arch width at the deciduous canines, size at 4 years with gain from 4 to 8 years (10); and dental arch width at the deciduous second molars, size at 4 years with gain from 4 to 8 years (18). In all instances the obtained relationships were low, statistically nonsignificant, or zero.

The second variety of velocity association investigated for skeletal nose height was the correlation between increase from 5 to 7 years of age and increase from 7 to 12 years of age. An answer was sought to the question, Can rate of growth in late childhood be predicted efficiently from rate of growth in early childhood? Again, analyses were made for both absolute and relative gains; the results are $r = 0.18$ and $r = 0.15$, respectively. In neither case can the null hypothesis be rejected; predictive possibilities are negligibly poor.

FINDINGS ON RELATION WITH OTHER VARIABLES

Velocity Correlations

Associations were investigated between gain in skeletal nose height during the period from 5 to 10 years of age, and gain in each of seven other dimensions over the same period. Six of the dimensions, together with references that describe the procedure employed in their measurement, are: head length (11), head width (13), upper face width (14), lower face width (21), hip width (17), and stature (19). The remaining dimension, subnasal height, was determined as the distance from the most anterior projection of the anterior nasal spine to menton. As in the previous section,

employing the Pearson r formula, both absolute and relative increments were correlated. Table 3 displays the coefficients obtained. Selected findings are:

1. It is not possible to reject the hypothesis of zero correlation between increase in nose height and increase in that portion of face height below the nose.³ Also, null hypotheses cannot be rejected for gain in nose height with gain in bigonial face width, and gain in bi-iliocrisal hip width.

2. The hypothesis that there is no relationship between increase in nose height and increase in head length can be rejected at the 5 per cent level of confidence. For gain in nose height with gain in stature, the null hypothesis can be rejected at the 1 per cent confidence level.

3. All of the increment relationships studied for the childhood span from 5 to 10 years of age are weak. This is shown by correlation coefficients ranging from low positive to zero.

TABLE 3

INCREASE IN SKELETAL NOSE HEIGHT BETWEEN 5 AND 10 YEARS OF AGE
(VARIABLE X) CORRELATED WITH INCREASE IN SEVEN OTHER
VARIABLES OVER THE SAME AGE PERIOD (VARIABLE Y)

Variable Y	N	r^*	p	$r^\dagger$	p
Head length	80	0.27	0.05	0.23	0.05
Head width	80	0.24	0.05	0.18	...
Subnasal height	66‡	-0.05	...	-0.05	...
Bigonial face width	80	0.12	...	0.12	...
Bizygomatic face width	80	0.25	0.05	0.19	...
Bi-iliocrisal hip width	80	0.17	...	0.21	...
Stature	80	0.34	0.01	0.35	0.01

* Coefficients for absolute (centimeter) increments.

† Coefficients for relative (percentage) increments.

‡ Increments for subnasal height were obtained only from roentgenograms showing occlusion of the upper and lower teeth at both ages.

Velocity Trend Comparisons

It has been stated in recent publications that velocity curves for facial dimensions show "a general circumpuberal rise" (20) and, with more explicit reference to the subject matter of this paper, that there is "a distinct acceleration of nose growth" at adolescence (27). The applicability of these statements to growth in skeletal nose height cannot be fully evaluated with data extending only to 14 years of age. There is the possibility, however, of supplementing the findings of the earlier section on magnitude trends

³ At age 5 years, roentgenograms satisfactory for deriving measures of the nasal and subnasal components of skeletal face height were available on 138 children. The coefficient obtained from correlating these measures is $r = 0.17$.

CHILD DEVELOPMENT

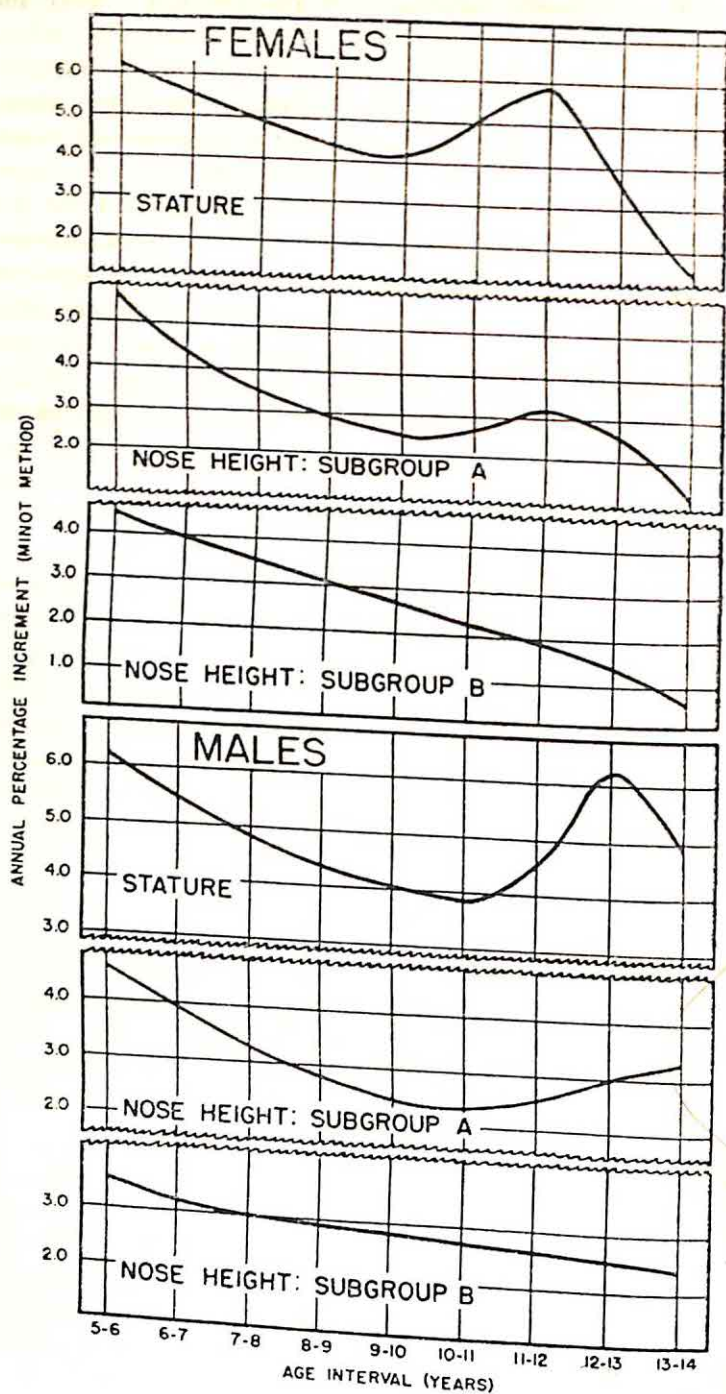


FIGURE 4—Smoothed velocity curves for selected stature groups and skeletal nose height subgroups (see text).

with some findings from comparing velocity trends for nose height and stature.

From among the 45 subjects on whom data were available to age 14 years, it was possible to draw two small groups that were homogeneous with respect to age at which the crest of the circumpuberal acceleration in stature was attained. One group consisted of seven girls cresting between 11 and 12 years; the other of seven boys cresting between 12 and 13 years.

On each member of both groups, annual percentage increments were computed for stature and skeletal nose height. The percentage increments for stature were tabulated in age-sex classes. Their subsequent analysis is indicated by stating that the first and fourth velocity trends shown in Figure 4 are drawn to means for these classes.

The percentage increments for nose height were first tabulated in age-sex classes and then subgrouped on the basis of whether in the period between the eighth and fourteenth years the consecutive values for individuals revealed presence (Subgroup A) or absence (Subgroup B) of acceleration. The second and fifth velocity trends in Figure 4 are drawn to age-sex means for the two girls and three boys of Subgroup A; the third and sixth trends to means for the five girls and four boys of Subgroup B.

Considering both sexes together, individual velocity trends for nine of the 14 children (more than 60 per cent) afford no evidence of a circumpuberal acceleration in nose height. The magnitude trends of these children are distributed as follows: concave, four; rectilinear, three; rectilinear-concave, two. Of the five children registering acceleration, only the two girls have similar velocity trends for nose height and stature. It follows that the magnitude trends of these girls are concave-convex-concave. The integral counterparts of the rising velocity trends for nose height on boys are concave-convex (two) and rectilinear-convex (one).

It is concluded that the findings from comparing velocity curves for stature and nose height call in question the generalization that "an adolescent spurt occurs" in "every muscular and skeletal dimension" of the face (27). The discrepancy between such breadth of generalization (5, 20) and the findings of the present study cannot tenably be resolved by recourse to time lag and truncation, i.e., by taking the view that for many of the individuals represented in Figure 4 a circumpuberal acceleration in nose height occurs after the fourteenth year. Nanda (20), from a comparative study of velocity curves for stature and five facial dimensions⁴ on 15 children examined repeatedly during childhood and adolescence, finds that where there are circumpuberal velocity increases in facial dimensions their timing is such that "the circumpuberal maximums in height and in the facial dimensions generally occur close to each other."

⁴ Not including nose height.

CHILD DEVELOPMENT

SUMMARY

The problems investigated center on individual growth in skeletal nose height during the postnatal period from age 5 years to age 14 years. Specific objectives relate in part to nose height *per se*, and in part to associations between growth in nose height and growth in other body dimensions.

Highly valid data for nose height are derived from roentgenograms of the face obtained at consecutive annual ages. These are supplemented with longitudinal data for subnasal height, widths of the upper and lower face, head length and width, hip width, and stature. The subjects are North American white children of northwest European ancestry. It is found:

1. For the age span from 5 to 12 years, there are three types of individual magnitude trend for nose height: an ascending curvilinear trend concave downward, an ascending rectilinear trend, and an ascending curvilinear trend concave downward in the first half of the period and convex downward in the last half.

2. Over the seven-year interval between ages 5 and 12 years, the increase in nose height for some children is twice that for other children. This range of variability holds with respect to both absolute and relative increments.

3. There is little, if any, relationship between the growth rate of nose height in early childhood (5 to 7 years) and its growth rate in late childhood (7 to 12 years). The obtained r 's, computed from centimeter and percentage gains, do not allow rejection of the null hypothesis.

4. For the childhood period from 5 to 10 years, growth in nose height is not highly related to growth in any of the other variables studied. Obtained associations with all seven traits lie in the zone between low positive and zero.

5. More than one-half of a small series of individual velocity trends for nose height, extending from one to three years beyond the age of the peak of the circumpuberal spurt in stature, give no indication of a circumpuberal spurt in nose height. The impact of this finding on current growth generalization is noted.

REFERENCES

1. ALLEN, I. Facial growth in children of five to eight years of age. *Hum. Biol.*, 1948, 20, 109-145.
2. ASHLEY-MONTAGU, M. F. The determination of the naso-frontal suture and the nasion in the living. *Radiology*, 1937, 28, 473-476.
3. BALDWIN, B. T. Physical growth and school progress. *U. S. Bur. Educ. Bull.*, 1914, No. 10 (Whole No. 581).

HOWARD V. MEREDITH

4. BOWDITCH, H. P. Comparative rate of growth in the two sexes. *Boston Med. & Surg. J.*, 1872, 10, 434-435. (See also, Report Brit. Assoc. Adv. Sci., 1880, 50, 140-147.)
5. DAVENPORT, C. B. Post-natal development of the human outer nose. *Proc. Amer. Philosophical Soc.*, 1939, 80, 175-355.
6. GOLDSTEIN, M. S. Changes in dimensions and form of the face and head with age. *Amer. J. phys. Anthropol.*, 1936, 22, 37-89.
7. GRAY, H., & AYRES, J. G. *Growth in private school children*. Chicago: Univer. Chicago Press, 1931.
8. HERSKOVITS, M. J. The anthropometry of the American Negro. *Columbia Univer. Contr. Anthropol.*, 1930, No. 11.
9. HIGLEY, L. B. Head positioner for scientific radiographic and photographic purposes. *Int. J. Orthodont.*, 1936, 22, 699-704.
10. HOLCOMB, A. E., & MEREDITH, H. V. Width of the dental arches at the deciduous canines in white children 4 to 8 years of age. *Growth*, 1956, 20, 159-177.
11. KNOTT, V. B. Physical measurements of young children. *Univer. Iowa Stud. Child Welf.*, 1941, 18, No. 3.
12. KROGMAN, W. M. Dental arch form and facial growth pattern in healthy children from prehistoric populations. *J. Amer. dent. Ass.*, 1938, 25, 1278-1289.
13. MEREDITH, H. V. Growth in head width during the first twelve years of life. *Pediatrics*, 1953, 12, 411-429.
14. MEREDITH, H. V. Growth in bizygomatic face breadth during childhood. *Growth*, 1954, 18, 111-134.
15. MEREDITH, H. V. A longitudinal study of change in size and form of the lower limbs on North American white schoolboys. *Growth*, 1955, 19, 89-106.
16. MEREDITH, H. V. Longitudinal anthropometric data in the study of individual growth. *Ann. New York Acad. Sci.*, 1955, 63, 510-527.
17. MEREDITH, H. V., & CARL, L. J. Individual growth in hip width: A study covering the age period from 5 to 9 years based upon serial data for 55 nonpathologic white children. *Child Developm.*, 1946, 17, 157-172.
18. MEREDITH, H. V., & HOPP, W. M. A longitudinal study of dental arch width at the deciduous second molars on children 4 to 8 years of age. *J. dent. Res.*, 1956, 35, 879-889.
19. MEREDITH, H. V., & MEREDITH, E. M. Annual increment norms for ten measures of physical growth in children four to eight years of age. *Child Developm.*, 1950, 21, 141-147.

CHILD DEVELOPMENT

20. NANDA, R. S. The rates of growth of several facial components measured from serial cephalometric roentgenograms. *Amer. J. Orthodont.*, 1955, 41, 658-673.
21. NEWMAN, K. J., & MEREDITH, H. V. Individual growth in skeletal bigonial diameter during the childhood period from 5 to 11 years of age. *Amer. J. Anat.*, 1956, 99, 157-187.
22. QUENOUILLE, M. H. *Associated measurements*. New York: Academic Press, 1952.
23. SCAMMON, R. E. The first serial study of human growth. *Amer. J. phys. Anthropol.*, 1927, 10, 329-336.
24. SHOLL, D. A. Regularities in growth curves, including rhythms and allometry. In E. J. Boell (Ed.), *Dynamics of growth processes*. Princeton: Princeton Univer. Press, 1954. Pp. 224-241.
25. SMYTH, C., & YOUNG, M. Facial growth in children. *Med. Res. Council*, Special Report Series, 1932, No. 171.
26. TANNER, J. M. The assessment of growth and development in children. *Arch. Dis. Child.*, 1952, 27, 10-33.
27. TANNER, J. M. *Growth at adolescence*. Springfield: Charles C. Thomas, 1955.
28. WYLIE, W. L. The relationship between ramus height, dental height, and overbite. *Amer. J. Orthodont.*, 1946, 32, 57-67.

THE INFLUENCE OF VIOLENCE IN THE MASS MEDIA UPON CHILDREN'S ROLE EXPECTATIONS¹

ALBERTA ENGVALL SIEGEL

Stanford University

With the arrival of television upon the social scene, there has been a renewal of widespread concern that the prevalence of violent themes in the dramatic presentations featured by the mass media may have harmful effects on the audience, especially the child audience. This concern had earlier focused on violent presentations in radio and in movies, and it has been expressed in countless articles and discussions in the popular press and even in the conduct of a legislative investigation of the entertainment industry (10).

Among professional workers, discussions of the possible effects of violence in the mass media have centered on the effects upon children's emotional development—upon anxiety, fears, aggressive motivation, etc. Recent experimental research has not confirmed hypotheses derived from this approach: Maccoby, Levin, and Selye (5) found no significant differences between frustrated and nonfrustrated children in their recall of aggressive movie content, and Siegel (8) found no significant differences in level of aggression or level of anxiety in children after they had seen an aggressive film and after they had seen a comparable nonaggressive one.

Much less attention, however, has been given to the possible influence of violence in the mass media upon what may be called children's *cognitive* development, which includes their conceptual learning of the realities of social life. It is with that possible influence that the present paper is concerned.

Among recent theoretical formulations concerning the variables of child behavior, Baldwin's has emphasized the importance of cognition as well as motivation in the instigation of behavior (1, pp. 112-130). Krech and Crutchfield have also stressed how importantly "cognitive contents" con-

¹ This study was conducted while the writer was the Josephine Glasgow Fellow of the American Association of University Women, to whom sincere gratitude is expressed. The writer is greatly indebted to Ellen Culbertson for participating in many aspects of the conduct of this study, to Professor R. D. Reifsnyder and his students at The Pennsylvania State University for producing the radio dramas, to Harold R. Glenn, Mabel Bresler, and Joan Wood of the Ferguson Township School in Centre County, Pennsylvania, for their generous administrative cooperation in the execution of the study, and to Jean Fuschillo, Joen Fagan, and Irma Shepherd for aid in the collection of data.

tribute to the direction of behavior (3, p. 248). And in classifying personality variables, McClelland includes in his basic triad "the individual's system of schemata," which encompasses "the individual's beliefs, frames of reference, major orientations, role perceptions, ideas, and values" (6, p. 242).

The purpose of the present experiment is to determine whether violence in the mass media influences the nature of that important aspect of personality to which the above cited terms refer. Specifically, the experiment concerns the influence of violence in the mass media on children's role expectations. Role expectations were chosen from the various sorts of cognitive contents enumerated by McClelland (above) not only because they are a sort which might well be expected to be influenced by the mass media but also because they are susceptible of operational definition in research with child subjects.

The concept of role is now a familiar one to behavioral scientists. Linton (4) has defined *status* as position in the social structure, and *role* as the behavior expected from any holder of a given status. Members of any society occupy several statuses—an age status, a sex status, an occupational status, etc.—and culturally assigned to each status is a pattern of expected behavior: a role.

An important function of the socialization process is to train the child to perform the behaviors (roles) associated with his various statuses. The socialized child in nursery school, for example, has learned the role behavior expected of him according to his sex, his age, perhaps his ordinal position in his family, and, to some extent, his socioeconomic status.

However, social life requires more. In addition to being familiar with the behavior requirements associated with one's own statuses, one must also know what behavior is associated with the various other social statuses, in order that one may interact with persons occupying these. To function effectively, the nursery school child must know what behaviors are to be expected from children of the opposite sex, from adults in the statuses of teachers, parents, "observers," "student teachers," etc. To a large extent, the adequacy of a child's social behavior will depend on the appropriateness of his "cognitive contents" concerning the role requirements of various social positions.

In accounting for children's behavior, we commonly refer to some postulated motive as explanatory. For example, an aggressive child may be said to have strong aggressive motivation. While the relevance and usefulness of such concepts cannot be denied, it is important to consider as well the child's cognitive learning in our attempts at explanation. If a child has learned to think that all spiders are dangerous, we need not attribute his destructive behavior toward a spider to any "aggressive motive." Moreover, aggression between children may sometimes be accounted for in terms of cognitive learning: either the child's concept of his own role may call for aggression, or perhaps his concept of the other child's role

leads him to expect aggression from the other and therefore arouses his own aggressive behavior in defensive offense.

This discussion is not to deny the importance of motivation in the instigation of behavior, but rather to call attention to the existence of other sources of instigation. These sources may well interact and modify each other, but it is important that conceptually we distinguish among them.

The term "role expectation" (7, p. 226) denotes a person's knowledge of what rights and obligations, actions and qualities, are appropriate to persons of a given social status. In the socialization process, role expectations are acquired not only through intentional instruction but also through incidental learning. Cultural agents may intentionally instruct a child concerning behaviors appropriate to his own statuses (age, sex, grade in school, socioeconomic level, religious affiliation, etc.) and occasionally concerning behaviors appropriate to the statuses of others. In addition, role expectations will be learned incidentally from a variety of sources. Incidental learning may be especially important in learning role behaviors appropriate to statuses held by others or to statuses which the child may himself hold at some future date. The sources of incidental learning about roles include not only living persons occupying various statuses (role models), but also literature, myths, jokes, stereotypes, etc., as well as explicit statements of social norms such as the Boy Scout Code (6, pp. 314-315).

To date, our information concerning the learning of role expectations comes from observation, recall, and other informal and uncontrolled sources. "Controlled experimentation on the learning of roles has not yet begun" (7, p. 227). The present study is an attempt at controlled experimentation in this field, and concerns one, possibly very potent, source for incidental learning of role expectations: dramatic presentations in the mass media.

The hypothesis under test is that role expectations held by children are modified by the content of dramatic presentations in the mass media. That is, the role behavior exhibited by various characters in dramatic presentations may be incorporated in the child's role expectations for actual persons occupying the statuses represented by the characters.

Many informal observations support this hypothesis. When Eastern children, especially those from urban areas, vacation in the West, they often expect ranch hands to behave like the cowboys they have seen so often in Westerns on TV and at the movies. Their fearful expectations of Indians reflect their faithful acceptance of the portrayals of Indians in these same sources. Closer to home, children commonly fear policemen, and this may in part be a result of their acceptance of the stereotyped portrayal of policemen in the mass media, wherein they appear as hostile and aggressive.

The general hypothesis will require many tests of its validity and generality. The extent to which children accept as socially accurate the role portrayals they see and hear in the mass media doubtless depends on such variables as: the extent and quality of their personal experiences with the role being portrayed; the consonance between the mass media portrayal

and other portrayals; the realism or irrationalism of the dramatic form used in the mass media presentations; the age, maturity, experience, etc. of the child audience.

The present study provides one specific test of the hypothesis by using a highly "realistic" form of dramatic presentation, using as subjects relatively immature and unsophisticated children, and focusing on a social status which is familiar to the child subjects but for which they have typically not developed highly structured expectations.

EXPERIMENTAL METHOD

Experimental confirmation of the hypothesis would be provided by a demonstration that different expectations of the same role will follow exposure to different dramatic presentations of that role in the mass media. Since this result is most likely in the case of role expectations which are not already definite or firmly structured, it was necessary to focus this experiment on a social role concerning which the subjects had no strong preconceptions. The subjects of the experiment were to be second graders, and therefore preliminary individual interviews were conducted with a small group of second graders (not members of the subject group) to determine their role expectations for a variety of social statuses held by men. (The pronounced sex differences in aggression in our culture led to the requirement that the roles be ones typically played by men. It was felt that any attempt to induce the subjects to include aggression in their role expectations for some women's status would have conflicted with the children's prior learning about sex-appropriate behavior.) From these informal discussions, the role of taxi-driver emerged as one with which the second graders were familiar, but concerning which they had no firm expectations or definite stereotype. Accordingly, the experiment used role expectations of taxi-drivers in the test of the hypothesis.

In brief, the experiment was conducted by exposing two equivalent groups of children to different series of dramatic presentations of the role of taxi-driver, on a medium of mass communication. Following these exposures, the children's role expectations of taxi-drivers were assessed in a different context. The hypothesis was that these expectations would differ in accordance with the contents of the dramatic presentations. The procedure is detailed below.

Subjects

All participants in the study were second graders in a nine-grade elementary school which is located in a small village and draws its students from that village and the surrounding rural countryside. The parents of the children were farmers, skilled and unskilled workmen, operators of small businesses, and institution employees.

In the school, the second graders were enrolled in two classrooms. These classes constituted the experimental (E) and control (C) groups of the present study. The acceptance of existing groupings for use in the experiment was justified on the ground that school administrators had assigned children to the two classes on a random basis; it permitted introducing the research procedures as part of the regular school routine. The descriptive data given below suggest that the E and C groups do not differ systematically, but rather are from a common population, as would be expected from the manner in which they were initially composed.

In the E classroom, 29 children were subjects of the study, having qualified by being in attendance for the dramatic presentations and the subsequent testing. In the C classroom, in which fewer second graders were enrolled, 15 children were subjects of the study.

The median age of the E subjects was 7-7, while that of the C subjects was 7-6. In both groups, girls were numerically preponderant: the E group included 18 girls and 11 boys, the C group, 9 girls and 6 boys.

The school records included intelligence test scores for 37 of the subjects. These scores were based on testing conducted nine months prior to the experiment, late in the subjects' first-grade year in school. Thurstone and Thurstone's Primary Mental Abilities test for ages 5 to 7 (9) was employed. This test yields IQ equivalents, normalized standard scores with mean 100 and standard deviation 16. For the E group, 23 first-grade IQs were available. These ranged from 93 to 118 with a mean of 107.2. For the C group, the 14 IQs which were available ranged from 93 to 119, with a mean of 107.6.

Both the ages and the IQs of the subjects suggest that the E and C groups were drawn from the same population. Acceptance of this proposition, which is based on statements of school administrators responsible for class composition as well as on the above data, permits us to attribute any observed post-experimental differences between the E and C groups to the experimental condition rather than to possible prior group differences.

Dramatic Presentations

Because exact control was required over the contents of the dramatic presentations on a medium of mass communication, those used in this study were written and produced especially for it. For reasons of feasibility in both production and presentation, radio dramas were used rather than movie or TV dramas.

Three radio dramas were written. Each was presented as one of a continuing series, "Taxi Driver." Each production was presented at the same time of day, and each used the same theme music, format, announcer, and introductory and concluding statements by the announcer. Thus, the productions had the usual characteristics of a series of radio productions. A taxi-driver was a central character in each of the three dramas, but the identity of the taxi-driver was different in each, for the purpose of the

study was to focus attention on the taxi-driver role itself, not on the possibly idiosyncratic characteristics of any particular person occupying that status. Each of the three productions employed devices to clarify the occupational characteristics of the taxi-driver role: the announcer discussed what a taxi-driver does, the character himself carried on brief conversations with his customers on this topic, and the story line incorporated the occupational characteristics of a taxi-driver. These devices were employed to insure that all of the listeners shared some common conception of what a taxi-driver does occupationally.

Each of the three dramas was written in two versions, an E and a C version. That initial part of each drama which set the scene, introduced the characters, and propounded the dramatic conflict was common to both versions. The two versions differed only in the *outcome* of the story: in the E version, the dramatic conflict was resolved by aggression and physical violence on the part of the taxi-driver, whereas in the C version another (nonaggressive) solution occurred.

The three dramas, each in two versions, were enacted and tape-recorded by students in an advanced class in radio production. For each drama, the participants in both the E and C versions were the same actors; in fact, the initial portions of the two versions, being common, were enacted but once and then this enactment was transferred to both the E and C tape recordings.

Content

Each of the six productions began with the theme music. The announcer's voice then broke in to identify the program, "Taxi Driver." When the music faded, the announcer made this opening statement:

Here is another story about a man who drives a taxi-cab. Every day, in every part of the country, many men make their living by driving taxi-cabs which carry other people. Their job is to drive the taxi. Taxi-drivers take people to work. They take passengers to meet trains. Taxi-drivers drive women downtown to go shopping. Sometimes they deliver children to their schools. A taxi-driver may drive a patient to see a doctor or to go to a hospital. A taxicab-driver has many adventures in his work. This program tells about adventures in the everyday lives of taxi-drivers.

After this standard statement, more of the theme music was played, and then the announcer introduced the day's episode. His introduction was followed directly by the enactment of the drama, which required about 10 minutes. At its conclusion, the theme music was heard again, and then the announcer returned with:

And that ends today's story of everyday adventures in the life of a taxi-driver. Listen in tomorrow at the same time for another thrilling story from the life of a taxi-driver!

After this statement, the theme music returned for a final time, completing the production.

Limitations of space preclude summarizing each of the three dramas here. Given below is a synopsis of the third drama, which will serve to exemplify the nature of all three. Its initial portion, common to both E and C versions, is given first, followed by the two alternative outcomes.

Drama III

Bill Miller is taking his woman taxi passenger home from Christmas shopping, and they discuss the problems of driving in icy weather. When she comments on the difficulty of his job, Miller launches into a lengthy boast about his perfect driving record, unbroken in years of fulltime taxi driving. The talk turns to another topic, but is interrupted by a loud crash, accompanied by shrieking brakes and screams of fright: the taxi and another car have collided. After determining that his passenger is not injured, Miller steps out to confront the driver of the other car.

E outcome: The two drivers blame each other for the accident. Miller becomes incensed at this accusation, citing his own perfect driving record. The other driver persists in blaming Miller, and suggests that he may petition for Miller's license to be revoked. At that, Miller threatens physical violence. When the other driver fails to yield in the face of Miller's abuse and threats, Miller makes good his threat to "beat you bloody." The passenger is shocked at the fight; when Miller returns to the cab, she expresses alarm that the other driver is lying motionless on the ground, hurt. Miller's only reply is, "I hope he *is* hurt." He turns his attention to the taxi. The passenger decides to walk to her destination, and Miller takes his cab to headquarters for repairs. There he boasts about having beaten up the man whose car hit his taxi.

C outcome: The two drivers begin by blaming each other for the accident, but Miller is mollified when he learns that the other driver is from California and is therefore unfamiliar with the techniques of driving over icy roads. When they turn their attention to their cars and to obtaining a report for insurance purposes, the passenger is consulted. In the course of the conversation, she recognizes the driver of the other car as her brother-in-law, with whom she had been out of touch for 20 years. Thoughts turn from the crash to the forthcoming reunion between her and her sister. Miller comments that he doesn't mind so much that his perfect driving record was broken, now that the crash has served to reunite a family.

Administration

The three dramas were presented to the children on three successive days (Monday, Tuesday, and Wednesday) at the same time-of-day. The purpose of presenting each drama at the same time of day was to preserve the illusion that the dramas were authentic radio broadcasts. The "radio" was a console which housed the tape player, and which was constructed to have the external features of a radio. It was wheeled in and out of the room for each presentation. The children were told that they were being given the opportunity to hear the programs so that the person who brought the radio to the room each day could determine whether these were the sort of programs enjoyed by second grade children.

Informal observation during the radio presentations indicated that the children enjoyed the dramas, typically became "involved" in them, and in no way questioned their authenticity.

CHILD DEVELOPMENT

Measure of Role Expectations

To assess the concept which each child held of the role of taxi-driver after hearing the radio dramas, a test was administered individually to each child. Testing was conducted on the two days following the final radio presentation, on Thursday and Friday of the same school week. The test, described below, was a completion test, orally administered on an individual basis, in which the child was asked to give completions to selected newspaper stories.

Rationale

The role expectations test was designed to elicit a child's *social* or *reality* conception of the role of taxi-driver, with special reference to his conception of the taxi-driver's resort to violence in his interpersonal relations in conflict situations. In certain structural features, the role expectations test bears a superficial similarity to projective tests of the story completion type (cf., 2). The nature and rationale of the role expectations test may be clarified by a discussion which contrasts it with tests of the projective type.

In a projective test, a primary aim is to tap the child's wishes, impulses, fantasies, motives, etc.; it is therefore necessary to minimize "contamination" of his responses by his conceptions of social realities, of what behavior is socially appropriate or standard. To accomplish this, certain techniques are typically used:

1. The test is introduced as a measure of imagination or creativity.
2. The instructions inform the subject that there are no "right" or "wrong" test responses.
3. In the stimulus material of the test, social reality features are minimized. In a picture test, for example, the status identities of the various characters may be unclear or ambiguous: the pictures are so constructed that the subject is not compelled by the stimulus material to identify any character as being of a particular age, sex, or other status position. Alternatively, social reality features may be avoided, as in some picture tests, by the use of animals rather than humans as story characters.

Each of these techniques is designed to minimize the social compulsion inherent in the stimulus material and therefore to maximize the subject's projection of his own impulses and wishes. The test material contains a minimum of features which would impose structure on the subject's response, and would therefore interfere with, or color, the operations of projection. The test instructions are formulated to serve the same end.

In the role expectations test used in this study, on the other hand, the focus is on the products of the child's social learning, and specifically on his understanding of what behavior is role-appropriate to a particular social position. Therefore, the technique aims to *minimize* the child's projection of his own wishes and needs onto his response. For this test, projected material can only be conceived as "contaminating." To minimize projec-

tion, several procedures are used which may be contrasted with those mentioned above:

1. The role expectations test is introduced as a measure of achievement, of how much the child knows.
2. The child is given to understand that there are "right" and "wrong" answers, and thus his own responses will be judged in relation to the correct ones.
3. Social reality aspects are *maximized* in the test materials. The stories to be completed are newspaper stories, and they are presented as being from a current edition of the local newspaper. The characters in the critical stories are identified according to their sex and other status positions, and the events of the stories are cast in a specific, concrete, and familiar local context.

Administration

The test was administered orally to each child in individual sessions. Testing was conducted in small private rooms in the school building. Two women with whom the children had no prior acquaintance administered the test. They followed a common procedure, having conferred in the development of standard techniques prior to the experiment.

Every effort was made to disassociate the testing from the radio dramas. The person who had brought the "radio" to the classrooms each day was not one of the testers, and was never seen at the school in either tester's company. Classroom teachers made no mention that the radio dramas and the testing had any relation to each other. The two testers did not appear at the school at any time before or during the presentation of the radio dramas. They made no reference to the dramas in their conversations with the children. If a child mentioned the radio dramas to the tester, she responded as if she had not heard of them before.

An assistant brought each child to the door of the testing room and introduced him to the tester. This procedure prevented the tester's knowing whether any child was from the E or C group.

When the child was seated, the tester said, "Jimmy, tell me what a newspaper is. What is a newspaper? What does it do?" The purpose of this opening inquiry was to determine whether the child understood that a newspaper reports events of the "real world," as distinct from presenting fiction. The child's response to this question was recorded verbatim, after which the tester entered a rating of either 0 or 1 in the child's protocol, indicating her judgment as to whether or not the child comprehended that a newspaper serves to report social reality.

In responses which received a rating of 1, it was common for the child to cite specific sorts of news which commonly appear in newspapers, and even to mention particular events which had been reported. One child mentioned that his sister's name was in the newspaper when she broke her arm. Another mentioned a news story reporting a fire in a neighbor's barn.

CHILD DEVELOPMENT

Many mentioned that automobile accidents are reported in newspapers. These responses are typical: "You get news in it. Like a wreck or something—you see it in the paper. Pictures too. You read in there what happened." "Tells what's coming up. Tells the weather. Wrecks and things." "It comes in the mail. You read it. It tells about things that happen: people get a baby, someone dies, someone gets hurt. My brother had his finger cut off and that was in the paper."

A child received a rating of 0 if he was unable to answer the inquiry, if his response merely described the physical characteristics of a newspaper, or if his response otherwise gave no indication that he understood a newspaper's reality reporting function. For example, one child said only, "You read it." Another said, "A paper, about fires and houses falling down," and another, "A thing that you read about things, like a wreck."

After the child had answered the initial question, the tester described the function of a newspaper, making this description sound like a confirmation, restatement, or extension of what the child had said. "Um-hum, a newspaper tells us what happens in the world, doesn't it? It reports the news: the things that happen to people. It tells about the activities of people in our community and in other places." This statement was followed by a pause to permit the child to make further remarks about newspapers.

Then the test was introduced.

Now, Jimmy, today I want to give you a little test. I want to find out how much you know about what is going on in the world. The way I'll do that is to find out about what you know about the news that is reported in the newspaper.

See these newspapers? They are copies of the *Centre Daily Times*, the newspaper for our area. I have chosen some news reports from these newspapers. I want to find out what you know about these reports. Here is what we will do. I will read to you the first part of a news report. Then I will stop, and you tell me the rest. You tell me how the situation came out. Now sometimes you may not know exactly, and then I want you to guess. If you do not know too much about the news story I am reading to you, then tell me how you would expect such a situation to work out.

This statement was made slowly to permit the child to respond as he chose, and statements which seemed unclear or confusing to the child were rephrased or repeated.

The tester had copies of two current issues of the local newspaper. On the inside pages of these, six "bogus" news stories, written for the purpose of this study, had been pasted. The tester was careful to hide these stories from the child as she turned the pages of the newspaper and read the stories. This procedure is not unusual in an achievement test, and neither tester was aware that any child suspected the authenticity of the news stories.

The six stories are given below in the order in which they were read. Subjects' completions to each were recorded verbatim.

The first two stories, concerned with events with which the children were familiar, were intended to give each child initial "success" experiences and also to confirm to him the reality-oriented nature of the test. These stories set the stage for the remaining four, which were entirely fictional and to which the child could not "know" the answers. In order to elicit acceptable responses to the first two stories, testers gave some prompting, if necessary, but their techniques were designed so that that child could think he had reached a correct answer independently. Responses to stories 1 and 2 were reinforced by the tester's enthusiasm and comments like "That's right," "Yes, it was cold, wasn't it?"

The testers gave no prompting or specific reinforcement for completions to stories 3 through 6. The initial instructions had encouraged the child to guess on stories for which he did not know the ending, and if necessary these instructions were repeated after story 3 had been read. To encourage the subjects, the testers used such techniques as these: nodding, saying "um-hmm," expectant pauses, questioning looks, holding pencil poised, etc.

Content

The six news stories used in this study were selected from a larger group which had been pretested with the pilot study group of children. Selection was based on the apparent comprehensibility and interest of the stories.

Story 1

There will be no school next Monday. All public schools will be closed, and school children will have a holiday. The reason that next Monday is a school holiday is . . . What is the reason?

The correct answer was that Monday was Lincoln's birthday. To aid the research, teachers in both classrooms had conducted discussions of that forthcoming holiday. However, some children suggested that the holiday was to celebrate Valentine's Day, Washington's birthday, Eisenhower's birthday, or Roosevelt's birthday, and these responses were accepted and praised.

Story 2

Winter weather has been visiting Centre County for the last few days. The roads have been slippery and drivers have found driving difficult. The sidewalks have been slippery too. In Pine Grove Mills (the name of the town in which the school is located), the weather has been . . . Now tell the news about the weather here.

These are typical responses: "Bad. Rainy. Snow." "Snowy and icy." "Kinda bad. We had some wrecks. People have been falling down and hurting themselves."

Story 3

Bill Smith was walking down Allen Street today, and he saw some money lying on the sidewalk. He walked over to pick it up. Just then Howard Brown stopped him. Brown said that he had lost the money earlier

CHILD DEVELOPMENT

in the day, and that he had been looking for it for a long time. Brown said that he was sure the money was his. Smith asked Brown to prove that this was the money he had lost. Brown could not prove it. So then Smith said that he was going to keep the money, because he had seen it first. . . . Now tell me what happened then.

Story 4

When John Henry got home from work today, he found a TV man there installing the new television set which the Henrys had ordered. As soon as the set was installed, the whole Henry family gathered around to watch television. They were excited about having their own TV set at last.

But neither Mr. Henry nor his wife could get the set to work. No matter what they tried, they couldn't get the TV set to work at all. Finally Mr. Henry called the TV company and told them to send the installation man back to the Henry house. The man came, and he couldn't make the set work either. When that happened, Mr. Henry told the TV man that he should take back the TV set and give them a good one instead. The man said that he couldn't do that. He said it was too late for Mr. Henry to get a different set. He also would not give back the money the Henrys had paid. The TV man told Mr. Henry, "You have already paid for that set and now you're stuck with it, even if it is no good!" . . . What happened then?

Story 5

Today was a busy day for taxi-drivers, for cold weather always brings extra business for taxi-cabs. But John O'Neal didn't earn as much money this morning as he might have. O'Neal, a taxi-driver for the local taxi company, was called to pick up a passenger at his home. O'Neal arrived at the address with his taxi at nine o'clock, and honked his horn to let the customer know that the taxi was ready.

The passenger did not come out, and O'Neal waited for him in his taxi for 15 minutes. Finally he went to the door of the house and knocked, in order to see what was keeping the passenger. The man came to the door and said he would be ready to go in just a minute. O'Neal went back to his taxi and waited there in the cold for another 20 minutes, which is a long time. Of course, he was losing money all this time, because he wasn't doing any business. Finally he went to the door and knocked again. This time the man came to the door and told the taxi-driver that he had decided to stay home and therefore he would not need the taxi. . . . What happened then?

Story 6

The cold and slippery weather brought bad luck today to Fred Horner, a taxi-cab driver. Up until today, Horner has had a perfect driving record: he never had an accident while driving his taxi-cab. But today there was so much ice on the streets that it was very hard to drive. So there was a crash between Horner's taxi-cab and another car. That crash broke Horner's perfect driving record. Both cars were damaged in the accident, but no one was hurt. After the crash, the taxi-driver got out of his taxi and . . . What happened then?

The reader will recognize that story 6 presents the same conflict as was presented in the initial portion of drama III. In this, story 6 differs from story 5, which is a "new" story—its plot does not duplicate or resemble the plots of any of the three radio dramas.

Scoring

Two methods of scoring were devised to permit two approaches to testing the research hypothesis. The first mode involved a comparison of the total amount of aggression included in a child's completions to stories 5 and 6 (the stories in which taxi-drivers figure) with the total amount included in his completions to stories 3 and 4. The second mode focused on the amount of aggression attributed to the taxi-driver in each child's completions to stories 5 and 6.

Difference in aggression. The first mode of scoring took into account children's individual differences in the amount of aggression they attribute to others. As would be expected, some children included aggressive incidents in every completion they gave, others included no aggression in any completion, and others ranged between these extremes. To eliminate the influence of such individual differences on scores used to test the hypothesis, a difference judgment was obtained.

Each child's completions to stories 3, 4, 5, and 6 were assembled. A judge read the four together, and compared the amount of aggression shown in completions 3 and 4 with the amount shown in completions 5 and 6. The latter stories concern taxi-drivers; the former do not. Any sort of aggression, from any source, was considered in this evaluation. The judge then assigned the child's protocol to one of two categories: *plus* or *zero-minus*. The *plus* category was for protocols whose completions to stories 5 and 6 contained more aggression than the completions to 3 and 4. The *zero-minus* category was for protocols whose completions to the two sets did not differ in aggression and for protocols in which completions 3 and 4 showed more aggression than completions 5 and 6.

Two judges made these categorizations independently without knowing which children were from the E and which from the C group. Their two sets of judgments showed 92.5 per cent agreement (total number of agreements / total number of judgments).

Taxi-driver aggression. Completions to stories 5 and 6 were also evaluated by another method. They were categorized according to the amount of aggression attributed to the taxi-driver.

The completions to the two stories were evaluated separately. For the judging, the E and C protocols were intermixed, and the judge had no knowledge of the group from which any protocol came. Judges were instructed to assess only the aggression attributed to the taxi-driver, not the total aggression attributed to all sources. Each completion was assigned to one of three categories, according to the amount of aggression attributed to the taxi-driver. A completion was assigned to the *high* category in cases where the taxi-driver fought, inflicted injury or pain, reported other persons to the police, got mad, yelled, was angry. A completion was assigned to the *intermediate* category if the taxi-driver showed some aggression (complained, accused, argued, etc.), or if the response implied that others

CHILD DEVELOPMENT

may fear a taxi-driver even though no aggression is ascribed to him, or if it is implied that opposing a taxi-driver is dangerous and may well bring trouble. A completion was assigned to the *none* category if no aggression was shown by the taxi-driver in it.

Each of the 44 protocols was scored independently by two judges to permit an assessment of the reliability of the scoring system. For completions to story 5, the judges were in complete agreement in categorizing 42 of the protocols, and differed by but one category in the other two cases. For completions to story 6, again the judges differed in only two cases, and then again by but one point. Thus, for each story, the judges showed better than 95 per cent agreement in their categorization of the completions.

Summary of Experimental Controls

Certain controls were included in the procedure to clarify the meaning of the outcome of the research. They have been mentioned in various contexts above, and are summarized here:

1. The E and C dramas were equivalent in format, acting personnel, length, initial portion, and in the manner in which they were introduced to the children. They differed in outcomes, in the attribution or nonattribution of aggression to the taxi-driver.
2. All evidence supported the assumption that the E and C subjects were from a common population, so that post-experimental differences need not be attributed to prior group differences but rather may be attributed to the treatment condition.
3. Every effort was made to disassociate the radio presentations from the subsequent testing.
4. Testers were not able systematically to identify the group (E or C) from which any child came when he was brought to the testing room.
5. Judges who scored the protocols had no extraneous information concerning the group (E or C) from which any protocol came.

RESULTS

The first step in the treatment of the data was to eliminate from consideration those children who gave evidence of not understanding the reality-reporting function of a newspaper, because these children's completions to the newspaper stories in the test could not be interpreted as reflections of their understanding of social reality. Children whose responses to the question "What is a newspaper?" were given a rating of 0 by their testers were dropped from the analysis. Four subjects received this rating. All four were boys, three from the C group and one from the E group. After they had been dropped from the subject group, the E group consisted of 10 boys and 18 girls while the C group consisted of 3 boys and 9 girls. Except in the ratio of boys to girls, the elimination of these subjects

did not render the E and C groups less equivalent. After the dropping of cases, the mean IQ in the E group was 108.5, in the C group, 108.4; the mean age in the E group was 7-7, in the C group, 7-6.

The mode of collecting the data permitted two approaches to testing the research hypothesis that the radio presentations would influence the children's role expectations of taxi-drivers, causing the E children to be more likely to attribute aggression to the taxi-driver role.

Difference in Aggression

The first test of the hypothesis utilized the difference judgments by which each child's completions to stories 5 and 6 (the taxi-driver news stories) were compared to his completions to stories 3 and 4. The prediction was that the proportion of subjects assigned to the two categories (*plus* and *zero-minus*) would differ in the E and C groups, with the E group being positively displaced relative to the C group. That is, the prediction was that, compared to the C group, the E children would have a higher incidence of aggression in completions 5 and 6 than in completions 3 and 4.

(If we could assume that stories 3 and 4 are *a priori* equivalent to stories 5 and 6 in the amount of aggression likely to be contained in children's completions to them, then we could simply predict that *plus* ratings would be more frequent than *zero-minus* ratings among the E children, because of the experimental treatment. Such an analysis would use each child as his own control and would not require the use of any separate control group. Since there is no good information to support the assumption that the two sets of stories are *a priori* equivalent, a comparable test is made by determining whether the E and C groups differ in their proportions of *plus* and *zero-minus* ratings.)

The data for this test are given in Table 1, which shows that the groups diverge in the predicted way. A chi-square test of these data shows the difference to be significant: with Yates correction for continuity, $\chi^2 = 3.29$, $p < .05$, one-tailed test. Only one-sixth of the C children had more aggression in their completion to the taxi-driver news stories than in

TABLE 1
DIFFERENCE IN AGGRESSION BETWEEN COMPLETIONS TO TAXI-DRIVER
NEWS STORIES AND TO NON-TAXI-DRIVER NEWS STORIES
BY E AND C SUBJECTS

	C Subjects	E Subjects	Total
<i>Difference in Aggression</i>			
Plus	2	15	17
Zero-Minus	10	13	23
Total	12	28	40

CHILD DEVELOPMENT

their completions to the others, while the majority of the E children had more. This is evidence for the effectiveness of the radio dramas in shaping the children's expectations of taxi-drivers.

Taxi-Driver Aggression

The second mode of analysis provides a more specific test of the research hypothesis, taking into account only that aggression attributed to taxi-drivers by the E and C children in their completions. This analysis is accomplished by a direct comparison of the E and C children's attribution of aggression to taxi-drivers in completions to stories 5 and 6. The data are the categorizations in terms of taxi-driver aggression: children's completions to these two newspaper stories were categorized as *high*, *intermediate*, or *none*, in terms of the amount of aggression attributed to the taxi-driver. The prediction was that the E children's completions would more frequently attribute aggression to taxi-drivers than would the C children's. The data for the test of this prediction are given in Tables 2 and 3.

TABLE 2
AGGRESSION ATTRIBUTED TO TAXI-DRIVER IN COMPLETIONS
TO NEWSPAPER STORY 5

	C Subjects	E Subjects	Total
<i>Aggression of Taxi-Driver</i>			
High	4	8	12
Intermediate	3	3	6
None	5	17	22
Total	12	28	40

Table 2 gives the data for newspaper story 5. In amount of aggression attributed to the taxi-driver in completions to this story, the two groups show no difference.

Table 3 gives the data for story 6. In amount of aggression attributed to the taxi-driver in completions to this story, the two groups differ significantly in the predicted direction: $\chi^2 = 8.40$, $p < .02$. (The small expected frequencies in some cells in Table 3 suggest that a supplementary test of these data be made. A more conservative analysis was performed by combining the "high" and "intermediate" cells. For the resulting 2×2 table, with Yates' correction for continuity, $\chi^2 = 2.81$, $p < .05$, one tailed test.) About half of the E children attributed *high* aggression to the taxi-driver in their completions to story 6, while no C child attributed *high* aggression to the taxi-driver and in fact two-thirds of them attributed *none* to him.

TABLE 3

AGGRESSION ATTRIBUTED TO TAXI-DRIVER IN COMPLETIONS
TO NEWSPAPER STORY 6

	<i>C Subjects</i>	<i>E Subjects</i>	<i>Total</i>
<i>Aggression of Taxi-Driver</i>			
High	0	13	13
Intermediate	4	6	10
None	8	9	17
Total	12	28	40

These data suggest that the main source of the difference in aggression attributed to taxi-drivers by E and C children is the difference in their completions to story 6. This is the story which paraphrased the initial part of radio drama III. Inspection of the contents of the completions to story 6 reveals many instances of literal carry-over from the contents of the radio drama to the contents of the completions. Of the 40 completions, 26 referred to the sex of the driver of the other car, and in every one the driver was identified as male, as was the case in drama III. Eight completions mentioned a passenger in the taxi, and each one identified her as female, as was the case in drama III. Two children mentioned the color of the car with which the taxi collided, and both identified it as blue, as it was in the radio drama. Many children said that the taxi-driver hit (fought with) the driver of the other car, of course, but only two mentioned any specific number of hits, and both gave that number as three, which it was in the radio drama. Six children's completions constituted more or less faithful recounts of the contents of the radio drama, either E or C, whichever they had heard. This one is perhaps the most complete and literal: "He checked his car. He hit the man three times. The lady in the car wanted to pick up the man, but he said 'No, he broke my perfect record.' The lady said she would walk to the store. He said 'Call me if you need help.' Then he drove to the station and reported 'A man came along and broke my perfect record.'" Four children, including two who were in the group just mentioned, made some overt reference to the radio dramas in their responses to story 6.

DISCUSSION

Interpretation of the Results

The interpretation of the results reported above rests on the assumption that the newspaper test has face validity as a measure of children's reality expectations. As the discussion of this test in an earlier section emphasizes,

all of the procedures of the testing were designed to establish this validity. Aimed toward the same end was the procedure of dropping from the analysis those children whose responses to the question "What is a newspaper?" did not indicate that they understood the reality-reporting function of a newspaper.

This is not to say that seven-year-olds distinguish between reality and fantasy with the same clarity as do adults, but rather it is to say that insofar as they do make this distinction they understand that newspaper contents are in the reality sphere. In point of fact, the hypothesis of this study is cogent precisely because the difference between reality and fantasy is so blurred and unclear for young children. The incorporation of many of the incidental features of drama III in many children's completions to story 6 illustrates this naivete.

Two modes of scoring the completions were used, and the statistical analysis of both sets of scores provides confirmation for the hypothesis that the radio presentations would influence the children's role expectations of taxi-drivers, causing the children who heard the dramas with endings *featuring* taxi-driver violence to be more likely to attribute aggression to the taxi-driver role. The first analysis demonstrated that the likelihood that more aggression would be included in the completions to taxi-driver news stories than in completions to other news stories was greater among the E than among the C children. The second analysis demonstrated that the E children attributed more aggression to the taxi-driver in their completions to one of the taxi-driver newspaper stories than did the C children.

Two interpretations may be attached to the finding that the E and C children did not differ in the amount of violence they attributed to the taxi-driver in their completions to story 5, but did differ, in the expected direction, in their completions to story 6.

Newspaper story 6 was the one which replicated the initial portion of drama III. Newspaper story 5 also concerned a frustration which came to a taxi-driver in the course of his work, but not a frustration which had been depicted in any of the radio dramas.

First, it may be that the children's learning from the three dramas was specific rather than general. That is, perhaps they learned only what is to be expected from taxi-drivers in the specific frustration situations depicted in the three dramas, rather than generalizing from this some inclusive formula like "taxi-drivers respond to frustration with physical aggression." Only when the reality situation duplicated the situation presented in the radio dramas were their expectations based on what had been presented in the dramas; hence, the group differences in response to story 6 but not to story 5.

Second, and alternatively, it may be that the children's learning from the three dramas *was* the generalization that taxi-drivers respond to frustration with aggression, but that the situation presented in newspaper story 5 was not perceived by them to represent frustration to the taxi-driver.

The frustration resulting from having to wait long and fruitlessly for a passenger may seem less severe than the frustrations depicted in the radio dramas, and therefore the formula "taxi-drivers respond to frustration with aggression" was not invoked as relevant.

The first interpretation could be tested by an experiment which gave the children greater opportunity to generalize from the radio dramas. One method would be to include many more radio dramas in the experiment, say 15 or 20, each depicting a different taxi-driver responding with physical aggression to a different frustration.

The second interpretation could be tested in an experiment in which other newspaper stories were used, stories in which the frustration to the taxi-driver is more evident to the children.

The present study, however, provides no specific evidence supporting either one or the other of these alternative interpretations.

Generality of the Hypothesis

The hypothesis that audiences generalize to social life from the contents of the mass media is a familiar one. It has long been recognized that the mass media have powerful "educative" capacities. Many have suggested, for example, that the stereotyped portrayals of Negroes in movie and television drama may serve to perpetuate and reinforce the American myth of ethnic differences, of "racial" superiority and inferiority. Those concerned with international understanding have frequently expressed concern over the representation of American life contained in Hollywood films exported abroad. A suspicion that dramatic portrayals in the movies are naively accepted as representations of reality underlies the fear that citizens of other nations may build a distorted picture of American life from their moviegoing experiences. It was this fear which the United States Ambassador to Italy, Clare Boothe Luce, invoked when she objected to the showing of the film "Blackboard Jungle" at the Venice Film Festival in 1955.

Similar reasoning underlies the argument that American children may acquire from the mass media a distorted conception of the realities of American life, particularly of the uses of violence in social life. This argument is also supported by the notion that children are less sophisticated than adults in distinguishing reality from fantasy.

We lack information concerning the conditions under which the carry-over occurs. Confirmation of the hypothesis of this study in one research context is not sufficient. The hypothesis requires many independent tests of its validity and generality. We need research directed to questions like those below.

Does the carry-over occur for roles with which children have had considerable personal experience as well as for those with which they have had little or none? For example, is the stereotyped portrayal of the "old maid schoolteacher" accepted as readily and unreservedly as the portrayal

of cowboys? Children have firsthand social experience with teachers, but usually not with cowboys.

How potent is a role portrayal in the mass media which runs counter to the role portrayals in other sources of role learning (jokes, stereotypes, literature, social experience, etc.)?

Are role expectations learned from unrealistic and highly fantasied dramas (for example, animated cartoons with animal characters, or stories of fairies, witches, pixies) as well as from realistic ones? We might guess that dramatic presentations in the "documentary" style have greater carry-over than do dramatizations of low realism, such as musical comedies, fairy stories, animated cartoons, etc.

How does the child's acceptance of mass media portrayals depend on his maturity? Young children, lacking experience and not always able clearly to distinguish reality from fantasy, may be more uncritical and unreserved in their acceptance of mass media portrayals than are more mature children.

A full evaluation of the influence of violence in the mass media upon children's reality expectations will require results from experimental approaches to answering the above questions.

SUMMARY

The hypothesis of the present study is that children's reality expectations are influenced by the contents of dramatic presentations in the mass media. More specifically, the hypothesis is that the aggression and violence they include in their role expectations of occupants of certain statuses is influenced by the attribution of aggression and violence to occupants of those statuses in dramatic characterizations in the mass media. Whereas discussion among professional workers concerning the influence of violence in the mass media upon children has centered on possible emotional effects, this hypothesis is addressed to possible cognitive effects.

To test the hypothesis, a series of radio dramas was presented to the children, after which their role expectations of the role held by the central character in the series were assessed by means of a newspaper test.

Subjects were second graders in two classrooms, E and C. The radio dramas were each enacted in two versions; the E children heard those in which the central character resolved the dramatic conflict by personal physical aggression, while the C children heard those in which another solution occurred to the same conflict. The E and C dramas shared a common opening, were enacted by the same personnel, and produced under the same format. They differed only in outcomes. On the days which followed the conclusion of the series, each child was tested individually for his role expectations of the status held by the central character in each drama, a taxi-driver. The test was a completion test, presented as an achievement test, in which the child was to give outcomes to various news

stories from current editions of the local newspaper. Subjects who did not understand the reality-reporting function of a newspaper were dropped from the analysis, for the children's responses to the newspaper test were taken as representing their reality expectations.

To permit two approaches to testing the hypothesis, two modes of scoring the completions were employed. Both were reliable: judges showed 92.5 per cent agreement on one scoring technique and 95 per cent agreement on the other.

The two sets of scores were used to test two predictions derived from the research hypothesis. The first prediction was that the E children would include more aggression, relative to the C children, in their completions to news stories featuring taxi-drivers than in their completions to others. This was confirmed. The second prediction was that in their completions to the news stories featuring taxi-drivers, the E children would attribute more aggression to the taxi-drivers than would the C children. This prediction was confirmed for one news story but not for the other. The news story for which it was confirmed was one which replicated the initial section of one of the radio dramas which the subjects had heard. Thus, in general, the data lend empirical support to the research hypothesis.

The discussion of the results includes an interpretation of the results summarized above and refers to other questions to which answers must be found in experimentation before an over-all assessment can be made of the influence of violence in the mass media on children's reality expectations.

REFERENCES

1. BALDWIN, A. L. *Behavior and development in childhood*. New York: Dryden, 1955.
2. FAUST, MARGARET S. Father-child relations in story completions. In Lois M. Stolz *et al.*, *Father relations of war-born children*. Stanford, Univer. Press, 1954. Pp. 244-264.
3. KRECH, D., & CRUTCHFIELD, R. S. *Theory and problems of social psychology*. New York: McGraw-Hill, 1948.
4. LINTON, R. *The cultural background of personality*. New York: Appleton-Century, 1945.
5. MACCOBY, ELEANOR E., LEVIN, H., & SELYA, B. M. The effects of emotional arousal on the retention of film content: A failure to replicate. *J. abnorm. soc. Psychol.*, 1956, 53, 373-374.
6. MCCLELLAND, D. C. *Personality*. New York: Dryden, 1951.
7. SARBIN, T. R. Role theory. In G. Lindzey (Ed.), *Handbook of social psychology*. Vol. I. *Theory and method*. Cambridge, Mass.: Addison-Wesley, 1954. Pp. 223-258.

CHILD DEVELOPMENT

8. SIEGEL, ALBERTA E. Film-mediated fantasy aggression and strength of aggressive drive. *Child Develpm.*, 1956, 27, 365-378.
9. THURSTONE, THELMA G., & THURSTONE, L. L. *Primary mental abilities test for ages 5 to 7*. Chicago: Science Research Associates, 1953.
10. *Television and juvenile delinquency: Report of the Committee on the Judiciary, United States Senate*. Washington: U. S. Government Printing Office, 1956.

GROUP CONFORMITY AND PERCEPTIONS OF PARENTS

PAUL HENRY MUSSEN

University of California, Berkeley

JEROME KAGAN

Fels Research Institute, Antioch College

A number of studies indicate that the tendency to yield to group opinion rather than to maintain independent judgment in experimental conformity situations is related to important personality variables (2, 3, 6). Thus, it has been found that college males who show a strong tendency to conform have a greater incidence of TAT dependency themes (*n* Succorance) than their less conforming peers, suggesting an essentially passive-dependent attitude toward adult sources of nurturance and support (6). Crutchfield's extensive studies of this problem reveal that, compared with those who are highly independent, extreme conformists are more authoritarian; more submissive, compliant, and accepting of authority; less tolerant, less socially active, and less able to accept responsibility; more anxious, over-controlled, and inhibited. Judging from their responses to items in personality inventories, conformists as a group are more intolerant of ambiguity and tend to emphasize conventional, socially approved values more than those who maintain independence in their judgments (2, 3). In interviews, they gave highly idealized descriptions of their parents, while independents offered a more balanced picture of praise and criticism. In general, extreme conformists have "rigid, externally sanctioned, and inconsistent moralistic attitudes" and "disturbed, dejected, and distrustful attitudes toward other people" (2, p. 195).

These data on the personality differences between conformists and independents suggest that early familial experiences may have an important influence on the adult's reaction to group pressure. More specifically, the conformist's submissiveness—reflected in his behavior in experimental conformity situations—may stem from his parents' early overly rigid and punitive treatment which prevented him from developing independence of thought and action. The independent's behavior, on the other hand, may be a consequent of less severe parental discipline. If this is true, and if the individual's perceptions of his parents, as revealed in fantasy productions, are correlated with early familial experiences, it may be hypothesized that extreme conformists would be more likely than independents to describe parents in TAT stories as harsh, restrictive, and punitive.

CHILD DEVELOPMENT

METHOD

The subjects, 27 male undergraduates, first wrote stories in response to eight TAT cards from the Murray series (1, 3BM, 6BM, 7BM, 12M, 13B, 14, and 18BM) and then participated in the conformity experiment designed by Asch (1). The procedure for this experiment followed Asch's instructions. The subject was led into a small room with four other men. Unknown to the experimental subject, these other four men were paid "stooges" cooperating with the experimenter. When the five men were seated, the experimenter gave the following instructions:

This is a task which involves the discrimination of lengths of lines. You see a pair of white cards in front. The card on the left shows a single line; the card on the right has three lines differing in length. They are numbered 1, 2, and 3 in order. One of the three lines on the card on the right is equal in length to the standard line on the card on the left. You will decide in each case which is the equal line. You will state your judgment in terms of the corresponding number, either one, two or three. There will be 12 such comparisons. Since the number of lines is few and the group small, I shall call upon each of you in turn to announce your judgment, which I shall record here on a prepared form. Please be as accurate as possible. Suppose we start at the right and proceed to the left.

The naive subject was always seated fourth from the right and was always the fourth person to call out his judgment. When all five subjects had stated aloud their judgments on one pair of cards, these two cards were removed and replaced by a new pair with new standards and comparison lines. This was repeated 12 times with 12 different pairs of cards. The measurements of the standard and comparison lines were identical with those used by Asch (1).

The four men cooperating with the experimenter had been previously instructed to announce correct answers on five trials (1, 2, 5, 8, and 11) and to announce incorrect answers on the other seven trials (3, 4, 6, 7, 9, 10, and 12). On each of these seven crucial trials, all four announced the same incorrect answer which then constituted a majority opinion.

Nine of the subjects, the high conformers, yielded to the incorrect group judgments on two or more of the crucial trials in the Asch situation, while 11 of them, the independents, never conformed to the incorrect group opinion. Seven subjects conformed only once and are not included in this study.

A clinical psychologist¹ analyzed the TAT protocols of the nine high conformers and the 11 independents, scoring each story for the presence or absence of parental punishment themes. Such a theme was considered present if parents in the story, individually or together, deprived, rejected, and/or physically or verbally punished their child. (Examples: "The boy

¹ The authors wish to acknowledge the assistance of Mr. Curtis Hardyk in this aspect of the study.

can't go out to play baseball because his mother is making him practice his violin lessons." "The mother is telling her son never to return home.")

The scoring was done "blind," i.e., the judge was not informed of the purpose of the study; hence, his TAT analysis could not have been biased by knowledge of the hypothesis. Since punishment was very precisely defined, scoring reliability was high. For 48 stories, scored independently by the judge and one of the authors (PM), the percentage of agreement was 93.

A *parental punishment* score was derived for each subject by simply counting the number of such themes in his stories.

RESULTS AND DISCUSSION

The two types of data obtained, conformity or nonconformity in the Asch situation and parental punishment scores, were used to test the hypothesis. For this group of subjects, parental punishment themes were relatively infrequent, the range being zero to three. Eight subjects had scores of zero, while six had scores of one. The distribution was therefore dichotomized into negative (no parental punishment themes) and positive (score of one or more) groups.

If the hypothesis is valid, a greater proportion of extreme conformists than of independents should have positive parental punishment scores. Table 1 shows the frequency of positive and negative scores among conformists and independents. Fisher's exact test for a 2×2 contingency table (4) was applied to these data and yielded the highly significant value of p between .025 and .01.

TABLE 1
DISTRIBUTION OF POSITIVE AND NEGATIVE PARENTAL PUNISHMENT
SCORES AMONG CONFORMISTS AND INDEPENDENTS

Group	PARENTAL PUNISHMENT SCORES	
	Positive	Negative
Conformists	8	1
Independents	4	7

Thus, the data support the hypothesis that conformity with the group is related to perceptions of parents, as reflected in fantasy productions, as harsh and punitive individuals. Assuming that these perceptions are the consequences of past experiences with parents rather than current views of authority in general, these findings suggest that the quality of early parent-child relationships may exert an important influence on the adult's conformity tendencies.

CHILD DEVELOPMENT

The data from Crutchfield's study, as well as from the present one, are consistent with the interpretation that conformity with the group may be a response which has been generalized from early-learned reactions toward feared parents. As Anna Freud has pointed out (5), a child may be able to reduce the anxieties stemming from parent-child relationships by *identification with the aggressor*, i.e., by imitating, and conforming with, the behavior of the parent he fears. Those who feel harshly treated by their parents probably have more conflicts with them and hence probably have more experiences in which conformity is rewarded, i.e., anxiety-reducing. Consequently, their conformity responses are likely to acquire relatively great habit strength and to be generalized to relations with other authorities, including the authority of group opinion. In this connection it is interesting to note that in Crutchfield's study conformists revealed marked attitudes of distrust toward others. Such attitudes may also be generalized responses stemming from original fear and distrust of the parents.

SUMMARY

The purpose of the study was to relate conformity tendencies to perceptions of the parents as shown in fantasy. Twenty-seven male college students wrote stories in response to eight TAT cards and then were individually observed in the Asch conformity situation.

A significantly greater proportion of extreme conformists than of independents perceived their parents as harsh, punitive, restrictive, and rejecting. These data, like those from previous studies, suggest that conformity tendencies are manifestations of basic personality structure and are influenced by early parent-child relations. The data may also be consistent with the interpretation that extreme conformity in the experimental situation represents a generalization from early-learned conformity behavior.

REFERENCES

1. ASCH, S. E. *Social psychology*. New York: Prentice-Hall, 1952.
2. CRUTCHFIELD, R. S. Conformity and character. *Amer. Psychol.*, 1955, 10, 191-198.
3. CRUTCHFIELD, R. S. Recent studies in conformity behavior. Paper given at Twelfth Regional Meeting of Veterans Administration Psychologists, California Academy of Sciences, San Francisco, November 17, 1955.
4. FISHER, R. A. *Statistical methods for research workers*. (7th Ed.) Edinburgh: Oliver & Boyd, 1938.
5. FREUD, ANNA. *The ego and the mechanisms of defense*. New York: International Universities Press, 1946.
6. KAGAN, J., & MUSSEN, P. Dependency themes on the TAT and group conformity. *J. consult. Psychol.*, 1956, 20, 29-32.

THE BEHAVIOR-INFERRED MOTIVATIONS OF LATE- AND EARLY-MATURING BOYS

PAUL HENRY MUSSEN *and* MARY COVER JONES

University of California, Berkeley

A number of studies indicate that the adolescent's physical status, mediated by the sociopsychological environment, may exert profound and lasting influences on his personality (4, 5, 6). In our culture, physically retarded adolescent boys are in a disadvantageous competitive position in athletic activities and are likely to be regarded and treated as immature by both adults and peers (5). Hence, they are more likely than their physically accelerated peers to be personally and socially maladjusted during adolescence. Analysis of the TAT responses of early- and late-maturing boys about 17 years of age, subjects in the Adolescent Growth Study (2, 3), revealed that significantly more of the latter group had negative self-concepts, profound feelings of rejection by others, strong affiliative needs (especially for heterosexual affiliation), prolonged dependency needs, and rebellious attitudes toward parents. The psychological picture of the physically accelerated boys, on the other hand, seemed to be much more favorable. As a group, they appeared to have acquired more self-confidence, were less dependent and in need of help, and were more capable of playing an adult male role in interpersonal relations. Contrary to what had been predicted, the two groups did not differ in needs for achievement and recognition, and there was some evidence that early-maturers had stronger aggressive needs than late-maturers (6).

The motives reflected in fantasy productions are presumably covert and not readily apparent in ordinary manifest behavior. It has been demonstrated that, in many cases, manifest behavior and fantasy may be independent or even negatively related (7, 8). Hence, inferences about motivation based on the individual's overt behavior may be quite different from those made on the basis of his fantasy productions.

In the present study, we shall examine the *behavior-inferred* ratings of motivation assigned to early- and late-maturing boys. Under Else Frenkel-Brunswik's direction, almost all subjects in the Adolescent Growth Study were rated on nine drives about a year after they had graduated from high school (1). The strength of each drive was inferred from behavior

observed in a wide variety of situations. "Intuitive judgments about motivational tendencies were obtained in the form of standardized ratings from a number of judges well acquainted with the adolescents" (1, p. 134).

The nine drives, selected from Murray's list of needs, were briefly described as follows:

1. *Drive for Autonomy.* Striving for independence and freedom; desire to be free from social ties, to shake off influence, coercion, and restraint; relatively little care for conventions and group ideology; tendency to act as one pleases.

2. *Drive for Social Ties, Social Acceptance.* Desire to be generally well-liked; to conform to custom, to join groups, to live sociably, to be accepted by a group in any form, to make contacts.

3. *Drive for Achievement.* Desire to attain a high standard of objective accomplishments; to increase self-regard by successful exercise of talent, to select hard tasks; high aspiration level. (Rating scale starts from "no desire to accomplish something outstanding" and ends with "excessive demands on himself.")

4. *Drive for Recognition.* Desire to excite praise and commendation, to demand respect, social approval and prestige, honors and fame.

5. *Drive for Abasement.* Tendency to self-depreciation, self-blame or belittlement; to submit passively to external forces, to accept injury, blame, criticism, punishment; tendency to become resigned to fate, to admit inferiority and defeat, to confess, to seek punishment and misfortune; masochistic tendency.

6. *Drive for Aggression.* Desire to deprive others by belittling, attacking, ridiculing, depreciating.

7. *Drive for Succorance.* Desire for support from outside; from people, institutions, or supernatural agencies.

8. *Drive for Control (Dominance).* Desire to control one's human environment, by suggestion, by persuasion or command.

9. *Drive for Escape.* Tendency to escape all unpleasant situations; to avoid blame, hardship, etc., to project own failures on others or on circumstances; to gain immediate pleasure with inability to postpone pleasure; use of fantasy, etc. (1, pp. 141-142).

"The rating scale was used as a means of summarizing what the judges 'privately' think, or what their hypothesis was, about each subject, or what they otherwise would have used only in writing up an interpretative case study of individuals whom they observed over a long period of time" (1, p. 144). "Though ultimately referred to observed behavior, reference to the underlying motivations was established as the result of a complex process of inference utilizing more subtle indirect cues together with gross features of behavior" (1, p. 261).

Since the drive ratings were found to be significantly correlated with many ratings of social responses and personality traits, Frenkel-Brunswick concluded that they "helped to organize the . . . data on overt behavior observed in social situations" (1, p. 261). "One drive variable [may circumscribe] a family of alternative [behavioral] manifestations unrelated to each other. The meaning of the drive concept emerges in terms of fam-

ilies of divergent manifestations held together dynamically or genotypically, though often not phenotypically" (1, p. 262).

In view of the interrelationships discovered, it may be assumed that these drive ratings refer to "a level of personality which stands behind the surface of overtly displayed social techniques" (1, p. 144), but which is probably not generally as remote from behavior as drives reflected in responses to projective tests. Comparison of late- and early-maturers' motivational ratings may, therefore, provide additional information about the factors underlying behavioral differences between the two groups. The objective meanings of these ratings may be derived from examination of the variables significantly correlated with them.

SUBJECTS AND PROCEDURE

The 34 subjects of this study, part of a normal sample of 90 boys who were members of the Adolescent Growth Study (2, 3), were drawn from the two extremes of a distribution of rate of physical maturing. Sixteen of them had been among the most consistently accelerated throughout the adolescent period; the other 18 had been among the most consistently retarded.¹

Three members of the Adolescent Growth Study staff who were well acquainted with the subjects over a period of years served as judges. They independently rated each subject on each of the nine drives listed above. In making the ratings, the judges were asked "to forget about manifest behavior and to group the children according to assumed motivation rather than according to similarities of displayed techniques" (1, p. 139).

Each drive was rated on a five-point scale, a rating of one representing the highest degree, a rating of five representing almost complete absence of the drive in question. In the present study, the three judges' ratings on each drive were averaged to obtain the individual's motivational rating.²

RESULTS

Frequency distributions of the ratings of all subjects were made for all the drives. Each distribution was then dichotomized at the point which most nearly enabled the placing of half of the 34 subjects above, and half of them below, the dividing point. Subjects having scores above this point

¹ The present sample includes all those subjects among the 19 most consistently accelerated and 19 most consistently retarded boys in the Adolescent Growth Study who were assigned motivational ratings. Twenty-nine of them were also in Jones and Bayley's (5) group and 31 in Mussen and Jones' (6).

² The interrater agreements on the nine drive ratings assigned to all the subjects of the Adolescent Growth Study are fully reported by Frenkel-Brunswik. The average reliabilities, corrected by means of the Spearman-Brown formula, were between .70 and .80, with the exceptions of those for Autonomy and Succorance which were .55 and .63 respectively (1, p. 100).

CHILD DEVELOPMENT

were considered *high* in this particular drive; those with scores below this point were considered *low* in it. Chi square tests were used to ascertain whether or not high ratings in certain drives were more characteristic of one group (late- or early-maturers) than of the other.

Table 1 shows the number of late- and early-maturers with high ratings in each of the drives, the chi square values obtained, and the levels of significance.

TABLE 1
NUMBER OF EARLY- AND LATE-MATURERS ASSIGNED HIGH
DRIVE RATINGS

Drive	NUMBER RATED HIGH		Chi Square	p
	Early-Maturers (N = 16)	Late-Maturers (N = 18)		
1. Autonomy	7	10	.47	ns
2. Social Ties	6	13	4.37	.02-.05
3. Achievement	10	9	.53	ns
4. Recognition	7	9	.15	ns
5. Abasement	9	8	.47	ns
6. Aggression	5	12	4.27	.02-.05
7. Succorance	5	9	1.30	ns
8. Control	9	9	.13	ns
9. Escape	9	11	.09	ns

As the table shows, the two groups were significantly different in two of the nine drives. A greater proportion of late- than of early-maturing boys were rated high in *Drive for Social Ties*; *Social Acceptance* and *Drive for Aggression*.

While these drives overlap to some extent, they are by no means identical. The correlation between them, based on the entire Adolescent Growth Study male population, was .48 (1). It should be noted, however, that the early- and late-maturing groups did *not* differ significantly in other drives (*Drives for Recognition, Control, and Escape*) that were even more highly correlated with *Drive for Aggression* and/or *Drive for Social Ties* in the entire sample of adolescent boys (1).

It is of course possible that the independent (physical status) and dependent (motivational ratings) variables were not completely independent, i.e., that in evaluating underlying drives the judges were influenced by their own impressions of the subject's rate of physical maturation. However, it seems unlikely that this factor had any significant effects, since the judges were specifically instructed to rate "assumed motivations" rather than more superficial characteristics. Moreover, if the judges had

been considerably influenced by the subjects' rates of maturation, the two groups might be expected to differ significantly on more than two of the nine drive ratings.

DISCUSSION

The finding that late-maturers appear to have relatively high motivation for social affiliation and social acceptance is entirely consistent with Jones and Bayley's data which showed that observers rated these boys as higher in sociability, social initiative and eagerness (5). Moreover, the TAT protocols of the late-maturers revealed that these boys have strong needs for affiliation, particularly with the opposite sex (6).

On the basis of evidence reported in earlier studies, it has been hypothesized that the late-maturers' emphasis on social activity and their social initiative and participation are largely of an attention-getting, compensatory nature (5, 6). Examination of the overt social behavior and personality trait variables correlated with *Drive for Social Acceptance* lends support to this hypothesis. For example, high ratings in this drive tended to be associated with high ratings in *energy output*, *interest in opposite sex*, *social participation*, *social self-confidence*, *social stimulus value*, *popularity*, *self-assertiveness* (bossiness), *exuberance*, *orientation toward the opposite sex*³ (1). These are characteristics which indicate expressiveness, energy, eagerness and sociability, a constellation labelled "overt social activity" by Frenkel-Brunswik. At the same time, however, high *Drive for Social Ties* was associated with many traits considered to be indicative of emotional maladjustment: *attention-seeking*, *affectation*, *dependence on approval*, *excitedness*, *tenseness*, *impulsiveness*, and *exploitiveness* (1). Hence, it may be inferred that, while late-maturers often reveal high motivation for social affiliation and a great deal of social activity, their social techniques are often childish and affected. Moreover, their high social drives may be based on general insecurity (reflected in *tenseness* and *impulsiveness*) and basic feelings of dependency (indicated by *dependence on others* and *exploitiveness*). The last conclusion is consistent with the finding that the TAT protocols of late-maturers revealed strong underlying dependency needs (6).

The relatively high aggressive motivation ratings assigned to the late-maturers may also be attributed, at least in part, to their basic insecurities, feelings of rejection, and intense dependency needs. Frenkel-Brunswik's data showed that *Drive for Aggression* had considerably higher correlations with emotionally maladjusted behavior than with "overt social activity." Among the manifestations of the latter, only four (*energy output*, *interest in opposite sex*, *social participation*, and *social self-confidence*) were significantly correlated with *Drive for Aggression*. On the other hand, 10 indi-

³ The correlations between the drive ratings and social behavior are fully reported in Frenkel-Brunswik's monograph (1, p. 186-7).

CHILD DEVELOPMENT

cators of emotional maladjustment (*attention-seeking, affectation, dependence on approval, excitedness, tenseness, impulsiveness, frequent mood swings, selfishness, irresponsibility, and exploitiveness*) were associated with high ratings in aggressive motivation. Apparently, the personal and social maladjustments of the late-maturers are often reflected in overt behavior *which trained observers* are likely to interpret as stemming from strong underlying aggressive motivations.

This finding is particularly interesting in view of the fact that late-maturers had relatively *low* fantasy (TAT) aggression scores, while the early-maturers had *high* scores in this variable (6). The present data do not enable us to specify the determinants of this discrepancy between behavior-inferred aggressive ratings and TAT aggression scores. Some of it may be due to the middle-class status of most of our subjects. It has been suggested that in this social class cultural prohibitions prevent the overt gratification of aggressive needs and thereby increase their intensity in the individual's fantasies (7, 8). If this is true, it may be hypothesized that the early-maturers, having identified with a mature male role, have incorporated these prohibitions and are less likely to behave aggressively. As a result, they are more likely to express their aggressive motivations in fantasy. On the other hand, the late-maturers, being defiant of authority (6), may not accept the middle-class standards of behavior so readily. Consequently, they may express aggression more freely and overtly and, therefore, may reveal less aggressive motivation in their fantasy productions.

It may be concluded that these group differences in motivational ratings add further support to the general hypothesis that the late-maturing adolescent is more likely to be personally and socially maladjusted than his early-maturing peer. Apparently in our culture, the physically retarded boy is more likely to encounter a sociopsychological environment which may have adverse effects on his personality development. The early-maturer's experiences, on the other hand, seem to be more conducive to good psychological adjustment.

SUMMARY

In this investigation, observers' ratings of 18 late-maturing and 16 early-maturing adolescent boys on nine drives were compared. The results showed that high drives for social acceptance and for aggression are more characteristic of the physically retarded than of the physically accelerated.

The meaning of group differences in these ratings may be clarified by examination of the social behavior and personality correlates of the two drives. These suggest that the late-maturer's high social drives may stem from feelings of insecurity and dependence and are often manifested in childish, affected, attention-getting social techniques. Moreover, high aggressive drives tended to be associated with social behavior and personality characteristics indicative of social and emotional maladjustment. Thus,

among the physically retarded, the strength of these drives may be partially attributable to underlying feelings of inadequacy, rejection, and dependence.

In general, the data of this investigation support the findings of earlier studies which showed that, among boys, physical retardation may have adverse effects on personality. Physical acceleration, on the other hand, may be conducive to better social and psychological adjustment.

REFERENCES

1. FRENKEL-BRUNSWIK, ELSE. Motivation and behavior. *Genet. Psychol. Monogr.*, 1942, 26, 121-265.
2. JONES, H. E. Observational methods in the study of individual development. *J. consult. Psychol.*, 1940, 4, 234-238.
3. JONES, H. E. *Development in adolescence*. New York: Appleton-Century, 1943.
4. JONES, M. C. The later careers of boys who were early- or late-maturing. *Child Develpm.*, 1957, 28, 113-128.
5. JONES, M. C., & BAYLEY, N. Physical maturing among boys as related to behavior. *J. educ. Psychol.*, 1950, 41, 129-148.
6. MUSSEN, P. H., & JONES, M. C. Self-conceptions, motivations, and interpersonal attitudes of late- and early-maturing boys. *Child Develpm.*, 1957, 28, 243-256.
7. MUSSEN, P. H., & NAYLOR, H. K. The relationship between overt and fantasy aggression. *J. abnorm. soc. Psychol.*, 1954, 49, 235-240.
8. SANFORD, R. N., ADKINS, N. M., MILLER, R. B., COBB, E. A., et al. Physique, personality, and scholarship: a cooperative study of school children. *Monogr. Soc. Res. Child Develpm.*, 1943, 8, No. 1.



FACTORS INFLUENCING INFORMAL GROUPS OF ADOLESCENTS

HAROLD R. PHELPS *and* JOHN E. HORROCKS

*Institute for Child Development and Family Life
The Ohio State University*

It has been the purpose of this study to investigate attitudes and activities that lead to the formation of informal adolescent groups. A large portion of adolescent social life is organized around small informal groups. Such groups are the proving ground for the teen-ager's widening perception of his social role and the exploration of techniques to implement that role. The informal group plays a much more important part in the life of the adolescent than do formally organized groups and may be thought of as a focal area of experience in the process of coming of age. Observation of informal adolescent groups leads to the conclusion that while the group is more or less ephemeral in its membership and continuity it does center around certain commonalities possessed by its members. Membership is usually exclusive in that some youth are excluded or are made to feel uncomfortable when participating in the group's activities. An understanding of adolescent behavior would be advanced by an analysis of the kinds of activities and attitudes important in the formation and perpetuation of such groups with particular reference to those things which contrast one group to another as well as to the commonalities which bind them all together as part of the larger adolescent age peer society.

Previous studies of group activities of youth have characteristically centered around a normative survey of activities participated in on the basis of such categories as age, grade, sex, intelligence, and socioeconomic status. Important as these categories have proved to be it would seem that more meaningful classifications could be developed from the interrelationships existing among data concerning the behavior of groups. Age, sex, and place of residence together with similar categories should be considered simply as other variables along with such facts as whether a subject's group plays ball, goes fishing together, rejects people who date, or makes use of a certain playground in their community. Analysis of activity and

CHILD DEVELOPMENT

attitude data against such major categories as sex rests on the assumption that a casual relationship exists between the categories and the data analyzed, with the major categories playing the role of independent variables. With the exception of age and sex there is a growing body of data which would indicate that the assumed causal relations between the categories and various activities and attitudes may be reciprocal (1). As an example, some persons may remain in a lower socioeconomic classification merely because they retain specific patterns of behavior and attitudes. There is also ample evidence to indicate that measured intelligence may be influenced by the opportunity for certain social experiences (2, 3). Thus, it would seem a more fruitful procedure to place age, grade, sex, socioeconomic status, etc., within the context of activities and attitudes involved in belonging to a group and allow them to find their places within categories on the basis of their relationship with all the data involved. If any variable emerges as a category or dominates a category, it has been validated. The present study has adapted such a procedure.

SUBJECTS

The 200 subjects for the present study were obtained from the 800 children enrolled in grades 7 through 12 in the public and parochial schools of a central Ohio community of approximately 9,500 persons. The 200 comprised a sample drawn randomly from among 726 children, stratified according to grade, who filled out the questionnaire on which this study was partly based and for whom comprehensive personal information was available. Most of the youth in the study lived in the city although a few commuted to school from the surrounding rural area.

PROCEDURE

All 800 children in grades 7 through 12 were administered a questionnaire entitled *Recreation in Your Community*. Of the 800 questionnaires given to students, 726 were returned. The questionnaire was omnibus in type. Six major areas of information were investigated as follows: (a) identification of the children by name, age, sex, grade in school, address, and other data of similar nature; (b) location and identification of informal adolescent groups by asking each subject to give names and other identifying characteristics of his favorite group or groups; (c) information concerning group attitudes, such as inclusiveness and exclusiveness, group ideals, and estimated group reactions toward samples of typical adolescent behavior; (d) measurement of the subject's feelings toward his own group and toward other groups with which he was acquainted; (e) naming of the activities and meeting places of the groups identified; (f) measurement of each subject's satisfaction with his groups and their activities.

In order to avoid the subjectivity of compiling items for a questionnaire by trying to anticipate possible responses that subjects might make, a series of both structured and open-ended interviews were conducted. Interviews about group membership and related matters were held with 100 subjects selected at random from the children in grades 7 through 12. An analysis of the interviews together with an extensive survey and field observation of the community in which the study was conducted made it possible to construct appropriate and information-producing items about informal adolescent groups and their activities in that community. A preliminary form of the questionnaire was designed and administered to a small group of subjects as a means of checking for ambiguity, interest appeal, and answerability of the questions included. It was felt that with subjects of such a wide range of ages and backgrounds it was particularly important to employ language capable of being understood by the youngest class of respondents.

In its final form the questionnaire¹ consisted of six sections named as follows: (a) What you would like to do for recreation, (b) Your group, (c) How do you feel about your group? (d) Whom you like in your group, (e) What your group does, and (f) Where your group goes.

Wherever possible check type questions were used. Considerable effort was made to make the questionnaire interesting to those who were asked to fill it out. Among other devices intended to achieve interest, the questionnaire was illustrated with line drawings. Every effort was made to refer to the respondent as an individual and to make him feel that the study was really interested in, and highly valued, his opinions. This was done, in part, by directing the questions to the individual, by using personal pronouns, and by heading each group of questions by personal reference statements such as "Your Group." With the cooperation of community authorities it was also possible to promise children who answered the questionnaire that recreational needs revealed by their answers would, whenever possible, be incorporated in the recreational programs of the schools and other community agencies.² Questionnaires were filled in and left in the local school offices in envelopes marked "confidential." Subjects were assured that their questionnaires would be available to no one in their community.

In addition to the information obtained by the questionnaire the following data on each subject were available: (a) socioeconomic status based on Warner's Index of Status Characteristics (4), and (b) intelligence as recorded in the local school records.

¹ A detailed discussion of the questionnaire and its construction will be presented in a forthcoming article by Horrocks and Bowlus.

² At the end of the study, through cooperation with the City Recreation Commission and the public schools, a number of the children's desires as revealed by the questionnaire were taken into consideration in planning the community's recreational program.

pit (.34). Most frequent time of meeting was on Sundays (.38) and at night (.28).

Activities appealing to adolescents scoring high on Factor I were playing cards (.26), listening to the radio (.42), watching television (.29), and having parties (.26). They liked to play baseball (.32). It did not seem to matter much to them if members of the group smoked (.40) or were not of the same age (.28), but they tended to prefer that members of the group not be colored (.28).

Young people influenced by Factor I were likely to be 17 years of age or older (.57) and to be no younger than 14 (.56). They were for the most part members of grades 11 and 12 (.53), and there was some tendency for seventh graders (.38) *not* to show the influence of the factor. There was some evidence that adolescents of IQ 111 or better (.28) whose fathers were in the upper occupations (.28) and whose families were of upper socioeconomic status (.26) were *not* likely to name a group influenced by the factor.

The satisfaction of heterosexual interests and a desire for peer-group activities not associated with the home were important pressures in this factor. None of the activities named was home-centered or home-influenced to any great degree. Associated with extra-home interests appeared a need for ranging about and going some distance from the community for social participation. It seems reasonable to assume that the forces underlying Factor I are a combination of pressures leading to the assumption of an adult role. Certainly, getting away from the influence of the home and ranging over territory farther and farther away from the home and community are important aspects of becoming an adult in the American culture. Of equal importance in becoming an adult are the physiological and social pressures which culminate in sexual mating.

It is of interest that lower socioeconomic status and lower intelligence were somewhat associated with Factor I. One expects adolescents whose parents have limited financial resources to assume the adult role sooner than those whose parents are more comfortably fixed. Previous studies have established that earlier emancipation from home, earlier assumption of the adult role, and greater sex activity are associated to some degree with lower socioeconomic status.

It is proposed that the present factor be described as a: *Pattern of pressures leading to assumption of the adult role, emancipation from the home, and satisfaction of heterosexual interests.*

Factor II

One variable stood out markedly in this factor. People who sometimes cheat to get by (.86) were not acceptable to groups influenced by Factor II. There were many behaviors which were thought to be inappropriate for group members. People who gamble (.57), smoke (.57), tell off-color jokes (.44), "park" (.44), "neck" (.43), ride motorcycles (.40), are noisy (.38),

swear (.32), and race cars (.31) were not acceptable. It was preferred that members of the group be persons of about the same age (.30), who have a good reputation (.36), nearly always agree (.34), and think good manners are important (.32). They thought that it was important for members to go to church (.30) but it was *not* important whether they went to churches other than their own (.30). Playing cards (.28) tended to be a rejected activity. There was a tendency to reject those who lived in the best part of town (.28).

There was much evidence that both home and school were forces in producing Factor II. Adolescents were preferred in the group who did not get in trouble with teachers (.57), did not quit school (.42), and were good students (.33). In relation to family matters, they expressed a preference for people who "stick-up" for their families (.48) and usually mind their parents (.28). The groups frequently met at places which were either part of the school or sponsored by the school. The assembly at high school (.54), noon movies at school (.52), and the high school stadium (.27) were favored places to go. They went to school athletic events (.34) together. One of the elementary school grounds (.42) located in a better part of town was a meeting place, and they patronized a recreation center (.31) sponsored by the school and located in one of the elementary school buildings.

Groups influenced by Factor II frequently went to the city swimming pool (.53) and liked to go swimming together (.28). They went together to movie theatres (.51), but tended to avoid the drive-in theatre (.31) so popular with adolescent groups identified with Factor I. Bicycle riding (.45) and hiking (.27) were favored physical activities, but it was likely that groups influenced by the factor would not play baseball (.25). They sometimes watched television (.26).

They patronized a hamburger "joint" (.40) in the business district, but tended to avoid a dairy bar (.39) also avoided by adolescent groups influenced by Factors III and V. One of the best restaurants in town (.26) most popular with adolescents of high socioeconomic status influenced by Factor V and avoided by those of low socioeconomic status reacting to the pressures of Factor IV was a meeting place for the Factor II group. They expressed a mild interest in sitting in restaurants and talking (.26).

Adolescents naming groups scoring high on Factor II were likely to be in the eighth grade (.38) and to come from families who obtained high, composite, socioeconomic ratings (.31). Their fathers tended to be employed in an upper socioeconomic occupation (.26).

With the high concentration of variables indicating disapproval for behavior and attitudes which are generally considered to be socially unacceptable, it may be hypothesized that Factor II has to do with a moralistic point of view or a need to conform to certain social standards. The fact that many of the activities were connected with school, that there was a moderately high loading on identification with persons and activities having upper socioeconomic status, and an emphasis upon family loyalty

CHILD DEVELOPMENT

suggests that the pressures behind this factor may be largely those induced by an upper socioeconomic home and school companions from homes of such status. One inconsistency was the appearance of a loading, however small, on the rejection variable for those who live in the best part of town. Such inconsistency might possibly be explained as a simple stereotype and expression of the belief that one should not overtly choose one's friends on the basis of their living in the best part of town. Although loadings on this factor were relatively low on variables having to do with church, there is some reason to believe that Factor II may be fostered by religious background since many of the values held by young people scoring high in this factor are characteristic of numerous religious groups.

Factor II is designated as follows: *Pattern of pressures assuming the form of a moral code approved by the school and upper socioeconomic home.*

Factor III

Very prominent in Factor III were variables having to do with going to business establishments where ice cream and other dairy bar products might be purchased. Going to three dairy parlors had loadings in the .60's. Two of the dairy bars had no provision for remaining after the purchase had been made. For some reason which the writers were unable to determine, one dairy bar was rejected with a negative loading of .33.

Noticeable were variables related to school life such as going to school affairs (.46), studying together (.52), going to the public library (.48), to school lunch room (.39), and assembly at the high school (.39).

There were many activities included in the factor. Those with highest loadings were having picnics (.61), listening to the radio (.59) and records (.57), bicycle riding (.57), making candy (.56), playing cards (.59), roller skating (.55), and playing volley ball (.52). Having parties, watching television, walking around town, swimming and tennis had loadings in the .40's. Only one activity was rejected and that was playing baseball with a negative loading of .32.

Groups scoring high on Factor III were likely to have seven or more adolescents in them (.35) and tended to be named more often by girls (.37). Young people naming groups high in this factor were likely to be in the eighth grade (.27), preferred the group they named (.28), and have fathers in the middle occupational classification (.26). Whether or not the person lived in the best part of town (.29) or wanted to quit school (.36) did not seem to matter.

It is very likely that this factor represents a cluster of activities around which some informal groups are organized. Most of the activities are ones which would generally be approved by teachers and parents. In contrast with Factor I all of the activities are ones which can be carried on within a relatively short radius of the home. In fact, all could be done within the small community in which the study was carried out.

Perhaps Factor III is the result of supervision exercised by the family forcing the adolescent to engage in activities which may be carried out not too far distant from home or under the surveillance of responsible adults. This would appear to be a logical conclusion particularly since girls who are supervised much more closely in our culture had a relatively large loading (.37) on this factor. It is therefore proposed that Factor III be named: *Non-emancipated, home, school, and community centered activity pattern.*

Factor IV

Adolescents preferring a group characterized by behaviors and characteristics associated with Factor IV were most likely to be of low socioeconomic status. Their families very often received a composite, socioeconomic rating (.38) placing them in the lowest category. They lived in the poorest housing (.40) located in the poorest sections of the city (.31). It was quite unlikely that any of their fathers were to be found in occupations rated among those in the highest category (.41) utilized in the present study. There was a trend for them to have an IQ on group tests of intelligence less than 89 (.25).

They appeared to make discriminations concerning group membership along socioeconomic lines. It was preferred that group members' fathers be laborers (.89) and they were likely to reject people who lived in the best part of town (.69) and had more money than most (.50). Persons who lived in the same neighborhood (.75) and were about the same age (.34) were preferred. They liked people who had high aspirations expressing a preference for those who wanted to go to college (.74) and worked after school (.60). Strength was important to them (.84).

It was important that group members go to church a lot (.50) and go to other churches (.85). What was meant by going to other churches is difficult to understand. Perhaps it was not particularly important what churches their group members attended, but that they go a lot to whatever church they preferred.

They thought that it was important that members of their group be in early at night (.56). Some value was placed on good looks (.29) and having a car (.25). People who were always kidding (.55) and nearly always agreed (.32) were liked. Colored young people were sometimes acceptable in their groups (.29). Frequently they made candy (.50) when they were together and went to an abandoned slate quarry (.43) located in a poorer section of the city. Roller skating (.41) was a popular activity, but there was a tendency for them not to go to movie theatres (.30) as a group.

One restaurant (.38) of very poor appearance, located in the business district, was popular. It was also popular with the adolescents of Factor I. Three dairy bars (.33, .30, .28) were sometimes patronized, one of which was popular with adolescents of Factors I and III, but avoided by those influenced by Factor V. One of the drive-in type was also popular with

people influenced by Factor I. The remaining dairy-bar was popular in both Factors I and III.

Some restaurants were avoided. One (.40) of very modern appearance located in the business district and specializing in fancy candies was not patronized by young people influenced by this factor but was found desirable by the young people of Factors I, III, and V. The restaurant (.34) which was possibly the finest in the city and very popular with the upper socioeconomic adolescents of Factor V was also avoided.

Without question Factor IV is the result of socioeconomic differences which are reflected in the nature of the activities and attitudes finding their places in this category. There appeared to be identification with their socioeconomic peers and a rejection of others in more fortunate circumstances. Perhaps they sensed the social and economic differences between themselves and other adolescents and were trying to make up for them by having higher ideals as manifested by their attitude toward church attendance and their somewhat unrealistic aspiration of going to college. Lack of financial means must not have been the sole determiner of the kinds of activities they participated in for they were able to patronize several restaurants and dairy bars. It was significant, however, that they avoided to some extent places popular with young people of upper socioeconomic status.

It is proposed that Factor IV be considered a: *Pattern of activities and social values deriving from a very low socioeconomic status.*

Factor V

Groups influenced by Factor V were more often named by girls (.41) who did *not* live in the poorer sections of the community (.48) or whose fathers were *not* engaged in low status occupations (.43). They were more likely to have fathers in the highest status occupations (.34) and to live in the best part of town (.32). As a matter of fact, several of the socioeconomic variables had loadings ranging from .26 to .48, rather clearly indicating that adolescents scoring high on Factor V tended to come from upper socioeconomic homes and *not* from the lowest group in the community. There were no significant loadings on the middle group.

Adolescents scoring high on Factor V named groups with seven or more in them (.56). The activities most favored were having slumber parties (.78), going to two higher quality restaurants downtown (.75, .73), and one about three-fourths mile outside the community (.58) which required some kind of transportation to reach it. In general, they liked to eat (.39) and talk (.34) in restaurants but were likely to shun certain establishments serving dairy products (.40, .30, .25) which were popular with groups scoring high in Factors I and III.

They liked to dance (.53), make candy (.41), listen to records (.40), have parties (.66) and picnics (.53), and date (.50). Physical activities engaged in were tennis (.60), swimming (.44), and baseball (.35), but they tended *not* to participate in basketball (.32), hunting (.30), and foot-

ball (.26). It was their preference that members of the group recognize the importance of good manners (.35) and like to date (.28). The groups named were likely to be of heterosexual composition (.29) and to some extent participate in the activities of riding in cars (.29), going out of town to have a good time (.27), and walking around town (.26). They were more likely to meet at night (.31) or in the afternoon (.25).

Many of their activities were related to the school. They studied together (.54), went to school affairs (.46), met in the hallways of the high school (.45), attended school athletics (.44), went to a school sponsored recreation center (.43), and met at the public library (.42). They preferred that their group members be of the type which would not quit school (.41). The adolescents naming this type of group were likely to be in grades 11 and 12 (.32) and to have IQ's above 111 (.32) and not below 89 (.34). Also, at school they attended movies together (.39) and met outside the school building (.35).

Dominant in this factor were socioeconomic variables. Many of the activities depended upon the adolescent's being a member of a family in average or better economic circumstances. There was a definite trend for adolescents from low socioeconomic level families *not* to designate this type of informal group as their favorite.

In Factor V there are evidences of a need to satisfy heterosexual interests and a need to participate in activities away from the home and community. However, many of the activities, such as having slumber parties, listening to records, having picnics, etc., suggest that approval of the family and ability of the parents to provide certain facilities were important influences. Worth noting is the fact that much activity involved the school.

It may appear somewhat incongruous that a factor with such a high loading on having slumber parties would also have a loading on the item of naming a heterosexual group. Upon inquiry it was found that it is not uncommon for slumber parties to include boys in the early stages and the morning after.

Apparently Factor V involved a somewhat complex set of social relationships characteristic of upper socioeconomic adolescents in the last two grades of high school. Girls appeared to be most sensitive to the pattern of pressures resulting in Factor V.

It is proposed that this factor be briefly described as follows: *Pattern of pressures leading to the assumption of an upper socioeconomic, quasi-adult social role.*

Factor VI

Social conformity and a concern for appearances are central to Factor VI. Adolescents' groups influenced by this factor rejected people who ride motorcycles (.61), "park" (.52), make wisecracks (.48), desire to quit school (.38), and are noisy (.35). They tended to reject other adolescents who race cars (.29), smoke (.29) and "neck" (.25). A strong

CHILD DEVELOPMENT

preference was expressed for persons who are good-looking (.55), are good students (.42), work after school (.39), and are of about the same age (.31). It was important that others in the group not be afraid of being different (.30). There was a tendency for them to like people who stick up for their family (.28), nearly always agree (.27), and dress neatly (.25).

Adolescents in groups scoring high on Factor VI went places together mostly associated with the school (.48, .29, .27) and a small dairy bar (.27). They participated in only two activities, volleyball (.39) and tennis (.26).

A concern for good appearance and the rejection of a noisy, show-off type of social role seems to typify Factor VI. Undoubtedly, pressures from the school and the home have much to do with the type of social conformity expressed by this factor. At first glance, the fact that a positive value was placed on "not afraid to be different" seems a contradiction. However, it is not uncommon for adults trying to produce conformity in adolescents to point out that if one wishes to amount to anything he must not be afraid of being different from other people.

It is therefore proposed that this factor be characterized as: *Pressures toward social conformity manifested by a concern for good appearances and rejection of a noisy, "show-off" type behavior.*

Factor VII

Groups scoring high on Factor VII frequented a water gym (.66) with picnicking facilities located about three miles outside the city and the city swimming pool (.26). They also liked to go to a partially abandoned amusement park (.61) at which a miniature golf course was still in operation. Restaurants (.54) were favorite places for congregating and they liked to sit and talk in them (.28). A dairy bar (.30) favored by adolescents scoring high on Factors I and III, but to some extent rejected by those scoring high on Factor V, was mildly popular. Groups of this type were most likely to be named by boys (.58).

As groups they participated in the sports of golf (.58), football (.57), basketball (.43) swimming (.34), tennis (.32), and wrestling (.25). They went fishing (.52), hiking (.47), and hunting (.44) together. They sometimes played pool (.47), but tended to reject the activity of making candy (.26).

There was some evidence that groups scoring high on this factor may have been influenced to some extent by the school. They preferred that their group members be popular in school (.60). A recreation center (.53) housed in a school was a popular meeting place and they sometimes met outside the high school building (.41). It is not clear whether they played together at the high school stadium (.43) or went there to view athletic events.

They liked members of their group to be good looking (.37), and to work after school (.30), but they did not care if they swore (.37), would fight (.29), or got in trouble with teachers (.27). Members were acceptable

who were always "kidding" (.28), who "necked" (.26), and sometimes cheated (.26). It was not necessary that they dress neatly (.26).

The fact that groups influenced by Factor VII were most frequently named by boys and many of the activities engaged in by the groups are usually considered to be primarily of interest to males suggests that this factor may be largely the result of pressures involved in playing a masculine role. Many of the activities required that the adolescent have the freedom to move about and freedom from adult supervision frequently denied girls in our culture.

It is interesting that the groups named by adolescents scoring high on this factor were not necessarily unisexual in their composition. As a matter of fact, there was a 50-50 chance that the groups named were made up of both boys and girls. This suggests that some females may have been involved in this factor who were perhaps having difficulty in establishing their roles as feminine members of society. Worth noting also is the fact that no socioeconomic variables were related significantly to Factor VII. If this factor is an expression of the male role, apparently socioeconomic status has much less meaning for boys than it does for girls.

Factor VII gave some evidence of being positively related to the school, but some activities and values deemed quite important by the adolescents influenced by Factor II had negative loadings on Factor VII. Apparently, the adolescent male is allowed to hold some values contrary to those held by the school and upper socioeconomic families. Or perhaps, this is a manifestation of his rebellion against the restriction of such a moral code.

It is proposed that Factor VII be considered: *Pattern of pressures and needs involved in playing a masculine role.*

Factor VIII

Some kind of conformity seemed to be important in this factor. It was apparently necessary that members of the group strive to please parents and teachers. They preferred that their friends be good students (.43), plan on going to college (.36), and stay out of trouble with their teachers (.38). It was important that group members mind their parents (.40) and be in early at night (.32).

What other people thought of them appeared to be important to young people scoring high in this category. They liked people who had a good reputation (.42) and believed that good manners are important (.32). A preference was expressed for people who were not afraid to be different (.43). Apparently they were willing to maintain these standards at the expense of sometimes being considered a little "queer." Or perhaps, adults in seeking to control them explained that individuals who wish to amount to anything can not be concerned about appearing to be different from the general run of people.

They avoided the slate pit (.34) and the old county fair grounds (.35) which was a large wood lot at the edge of town. It was unlikely that they

would make candy (.31) when they got together. Their groups met at school movies (.35) and outside the high school building (.30), but to some extent avoided the school cafeteria (.26). Adolescents most frequently naming this type of group were in the ninth and tenth grades (.34).

It is somewhat difficult to interpret Factor VIII; however, it seems fairly apparent that adolescents under the influence of this factor must have been convinced of the importance of pleasing their parents and certain teachers. On the other hand, one can hardly resist conjecturing what the effect of living by such a code might have had on peer relationships, particularly at the ninth and tenth grade level. Could the adolescents influenced by this factor have been isolates because their standards made it difficult for them to establish satisfactory peer relationships? Only further, more definitive research can answer this question. In the meantime, it is tentatively suggested that Factor VIII be considered as: *Pattern of pressures resulting from adult domination and lack of emancipation from the home.*

Factor IX

Good appearances seemed very important in Factor IX. A strong preference was expressed for people who dress neatly (.68), have a good reputation (.62) and think manners to be important (.56). Frequent church attendance (.33) was desirable. Persons were likely to be rejected from their groups who told off-color jokes (.57), swore (.49), made wisecracks (.32), would fight (.49), or would gamble (.47). They preferred that members of their group be white (.66) and were likely to reject colored young people (.56). A high value was placed on minding parents (.61), and there was a tendency to prefer people who got in early at night (.27). It was important to avoid appearing as if one always had to win (.59) presumably in various kinds of games or discussions. Racing cars (.40) was frowned on.

Factor IX almost took on the appearance of being an expression of a feminine role. Perhaps it is an expression of a particular type of female role. Many of the activities important in Factor VII which appeared to be dominated by masculine interests were rejected. Adolescents of the present factor did not wrestle (.63) or play basketball (.61) and football (.55). Hunting (.53) and fishing (.45) were not at all popular. Playing pool (.53) was frowned on. Baseball (.29) and roller-skating (.27) were the only physical activities in which they engaged to any extent. They liked to go on picnics (.26).

Young people naming groups rated high on this factor came from families receiving a composite socioeconomic rating in the middle group (.40) for the community. Their fathers tended to be employed in occupations falling in the middle (.26) of the classification employed in the present study. Seventh graders (.37) were not likely to be found among those naming groups high in Factor IX.

Apparently the questionnaire failed to include any places that groups of this kind liked to go; however, they appeared to avoid some which were

listed. The playground of an elementary school (.49) located in the better part of town, school movies (.41), and the cafeteria at high school (.26) were not acceptable meeting places. The slate pit (.48) and the old fair grounds (.33) were not approved places to be seen. A dairy bar (.26) popular with adolescents of Factors I and III and avoided by those of V was to some extent avoided by young people scoring high on Factor IX.

The moving force behind Factor IX appeared to be a need for approval and status. Apparently this need found its expression in maintaining the very best of appearances, having good intentions, associating with the right people and avoiding activities which might mark one as an uncouth or perhaps unladylike person. Since groups scoring high on the present factor were most frequently named by middle-class girls, it is perhaps to some extent the manifestation of a middle class, feminine role. Apparently it does not become an issue until after the seventh grade.

Factor IX may suggest that the most frequent means for the middle class girl to climb the socioeconomic ladder is through a program of appearing to be highly virtuous, having the very best of intentions, and avoiding any activity which would detract from her femininity. On the other hand, such a pattern of activities, attitudes, and values might be an attempt to maintain status by appearing to be as different as possible from those in a lower socioeconomic classification.

It is significant to note that the school does not play a role to any extent in this factor, in spite of the fact that many of the values related positively to this factor were important in Factor II which was so closely allied to the school. There was, however, some indication that the home plays a part in the formation of the factor.

It is proposed that Factor IX be characterized as a: *Need for approval and status growing out of pressures applied by the middle class family.*

Factor X

After a thorough consideration of the variables which emerged in Factor X the meaning was so unclear that it did not seem wise to attempt an interpretation. It was the last and weakest factor to emerge from the analysis.

DISCUSSION AND CONCLUSIONS

Many of the variables in the present study had loadings on more than one factor. While an attempt is made in multiple factor analysis by means of rotation to maximize loadings on one factor and minimize them on all other factors, there are limits imposed by the very nature of the data under consideration. This is true because some activities and attitudes have varied meanings to adolescents and are used by them in different ways. As an example, dancing, with a loading on Factor I, appeared to be a means of assuming the adult role and to a considerable degree a means of satis-

fying heterosexual interests. In Factors III and V it appeared to take on the meaning of an activity by means of which relatively uncomplicated social needs are satisfied. Making candy in Factor III appeared to be an activity which is done around home and which in all probability requires the cooperation of parents. In Factor IV it appeared to be one of the few activities which can be done at the home of adolescents who are of very low socioeconomic status, while in Factors III and V making candy was one of many activities characteristic of upper socioeconomic social groups. In Factors III, V, and VII, making candy appeared to be a more feminine than masculine activity. In Factor VIII it took on the possible meaning of an activity which is refused children in a family exercising undue control. While the analyses just made of the two activities may be subject to a great deal of error, they are examples of the kinds of hypotheses which might be proposed on the basis of the present type of study.

There was evidence in the present study to indicate that socioeconomic status of the family is an important factor in connection with attitudes and activities of adolescents. It does not, however, stand out as a dominating factor. How much effect the family's socioeconomic status has upon the behavior of adolescents can not be determined from the present study. However, it does appear that a very important difference between the socioeconomic levels appears in the degree and quality of emancipation from adult control allowed the adolescent. Emancipation is likely to depend upon many factors. In all probability the role of an adult will be forced on those adolescents of lower socioeconomic status sooner because of the requirements of having to support one's self at an earlier age. But, aside from financial necessity, the way in which emancipation from adult control takes place is probably highly conditioned by standards and values depending upon culture, religion, philosophy, and education. It would seem that more important than socioeconomic status itself is the quality and degree of emancipation from adult control which appeared to be a major issue in factors not loaded to any degree on socioeconomic status.

Adolescent girls tended to mention groups that were typical of factors showing a considerable amount of association with socioeconomic status. This fact may indicate that girls tend to make distinctions along socioeconomic lines more often than do boys. Or perhaps it may be accounted for by the fact that girls tend to be under more strict adult supervision. Closer supervision and closer home ties may produce in girls a greater consciousness of status and values held by the family than is found in boys. Noticeable is a tendency for the upper socioeconomic factors to be associated with activities centered in the school. This suggests that the school tends to encourage attitudes and standards that are characteristically upper socioeconomic in nature. "Upper socioeconomic" in this small Ohio city probably means what one would ordinarily speak of as "middle class" for the United States as a whole. In this city there was a very small "upper" class insofar as one existed at all. In the socioeconomic typing there were

no houses that could be rated "1," using Warner's method. It was most difficult to make clear-cut distinctions as to the quality of the various neighborhoods.

Chronological age had a high loading on Factor I. It had barely significant loadings on Factors VII and X, but because they were below the cutting point of $\pm .25$ they were not used in the interpretation of the two factors. The fact that chronological age was associated to any degree with only one factor out of 10 suggests that it is by no means a dominant factor in adolescent behavior and attitudes. On the other hand, grade position in school was loaded significantly on several of the factors. Of the two, then, grade position is probably more useful in predicting adolescent group behavior. This may be due to the fact that roles are assigned adolescents which are characteristic for grade levels. In other words, it is likely that social maturity is gauged by the layman and peer groups more by school grade than by chronological age.

Intelligence tended to be related to the same factors upon which socioeconomic measures were loaded, thus bearing out the considerable amount of research that shows a positive relationship between intelligence as measured by the traditional mental test and socioeconomic status. There is no evidence in the present study to throw light on the nature of the relationship between measured intelligence and socioeconomic status.

There was a tendency for factors loaded on upper socioeconomic status to be loaded also on the variable "seven or more in a group," indicating perhaps that adolescents of higher socioeconomic status tended to associate with larger peer groups than did the adolescents of lower socioeconomic status. Apparently, adolescents of high socioeconomic level tended to make more social contacts and to be involved in more interpersonal relationships than those of a lower socioeconomic level. This fact may be of significance in the area of social skills. The upper socioeconomic status adolescent has much greater opportunity to practice the skills of social intercourse. As a result, he is likely to have a great advantage over his peers of lower socioeconomic status when it comes to effectively relating himself to others in the competitive situation which is characteristic of his social milieu.

While the naming of factors in the preceding section of the present study is in a real sense a statement of conclusions, it seems worthwhile to set forth several generalizations which grew out of the study as a whole.

1. The present study has demonstrated that in the area of psychosocial research multiple factor analysis can be a useful statistical method in the development of hypotheses and categories from empirical data all of which have had equal opportunity to influence the final outcome.

2. While socioeconomic status is significantly related to many areas of adolescent informal group attitudes and activities, it does not appear to be the dominant factor.

3. In her informal group activities and attitudes the adolescent girl is influenced to a greater degree by socioeconomic status than are boys.

CHILD DEVELOPMENT

4. Identical activities and attitudes may be used by adolescents to satisfy widely divergent needs.

5. Degree of emancipation from adult control appears to be a most important influence in the formation of patterns of informal adolescent group activities and attitudes.

6. The public school as a social institution tends to re-enforce upper socioeconomic values and is used to a considerably greater extent to satisfy social needs by upper-socioeconomic young people than by those of lower status.

SUMMARY

In the present study the area of adolescence was considered through the investigation of adolescents' reports concerning activities and standards of social acceptability of informal groups with which they claimed most frequent association.

A small city in Ohio was selected in which to conduct the research project. A questionnaire was developed and administered to students in grades 7 through 12 of the public and parochial schools. In addition to data gathered in the questionnaire, intelligence quotients and socioeconomic ratings were obtained for each subject.

An adaptation of multiple factor analysis was selected as the statistical tool best suited for analyzing the data. Ten factors emerged from the analysis, nine of which were tentatively identified.

REFERENCES

1. ANDERSON, H. H., & ANDERSON, GLADYS L. Social development. In L. Carmichael (Ed.), *Manual of child psychology*. (2nd Ed.) New York: Wiley, 1954. Pp. 1162-1215.
2. JONES, H. E. The environment and mental development. In L. Carmichael (Ed.), *Manual of child psychology*. (2nd Ed.) New York: Wiley, 1954. Pp. 631-696.
3. McCANDLESS, B. Environment and intelligence. *Amer. J. ment Def.*, 1952, 56, 674-691.
4. WARNER, W. L., & MEEKER, MARCIA. *Social class in America: A manual of procedure for the measurement of social status*. Chicago: Science Research Associates, 1949.
5. WHERRY, R. J., & GAYLORD, R. H. The concept of test and item reliability in relation to factor pattern. *Psychometrika*, 1943, 8, 247-264.

GENERAL INFORMATION OF KINDERGARTEN CHILDREN: A COMPARISON WITH THE PROBST STUDY AFTER 26 YEARS

MILDRED C. TEMPLIN

Institute of Child Welfare, University of Minnesota

Many descriptive studies of child behavior were carried on in the 1920's and the 1930's. Since that time the environment of young children has changed greatly, and the results of recent but comparable studies of behavior are both similar to and different from those carried on 20 or more years ago. Since the range of children's information should be particularly sensitive to the differential impact of a changed environment, the author repeated in 1954-1955 the study carried on by Probst in 1928 with kindergarten children (2, 3).

PROCEDURE

Sample

In the present study the number, sex, age and socioeconomic distribution of the subjects in the major sample are the same as in the Probst study. Both studies included 100 children, 50 boys and 50 girls between five years, four months, and six years, chosen to be representative of the total population of Minneapolis according to their father's occupations as classified on the Minnesota Scale of Paternal Occupations. In both studies the 50 children in classes I, II, and III (professional, semi-professional and managerial, clerical and skilled trades) constituted the upper socioeconomic group, and the 50 in classes V, VI, and VII (minor clerical jobs, semi-skilled trades, slightly skilled and unskilled labor), the lower socioeconomic status group. Probst's subjects were selected from seven Minneapolis Public schools and those in this study from 11 such schools in neighborhoods appropriate to their fathers' occupations. Although Probst had Detroit Intelligence Test scores available, in the present study no such data were available and no intelligence test was given. An indirect control of intelligence is achieved in the selection of the sample.

The children in this study were questioned during their first semester in kindergarten. Since Probst questioned her subjects during the second

CHILD DEVELOPMENT

semester, an additional 24 kindergarten children in socioeconomic class V enrolled in the same schools as the class V subjects in the major sample were interviewed the following May.

Questionnaire

The Probst questionnaire samples 11 categories of knowledge: I. Local points of interest; II. Time and number; III. Current topics and history; IV. Natural phenomena; V. Literature and music; VI. Animals, birds and insects; VII. Plants and flowers; VIII. Occupations and industries; IX. Household arts; X. Simple mechanics; XI. Games and amusements. Each form, A and B, was approximately equal in difficulty and contained 66 questions, six in each category of knowledge.

All the Probst questions were used in the present study so that specific comparisons are possible.¹ However, 21 new questions were added: 16 replacing questions judged out-dated in categories III, IX, X, and XI, and five parallel questions sampling current information in category III. These new questions together with the number of the Probst items they may replace appear in Table 4.

In both studies the children were questioned individually: by Probst in two separate sessions within a seven-day period, in this study in a single session. Probst reported a mean of 17 minutes for the administration of Form A and of 19 minutes for Form B, while the mean time for the single interview was approximately 30 minutes.

Probst's rules and principles were followed in scoring the original and the new questions. Some specific responses not credited in the earlier study might be in this one. For example, in response to "What is the Great Northern?" not only references to the train, trains, or the depot but also references to the Great Northern Markets (a local chain) were credited. The score in the Probst study was the number of correct responses. In this study the score on the original Probst questions is referred to as the "regular" score. The score achieved by replacing certain questions with the new substitute questions is designated the "substitute" score. For both regular and substitute scores the maximum value is 66 for each form and 132 for the entire questionnaire.

Analysis of Responses

Probst reports the percentage of correct responses given by boys, girls, upper and lower socioeconomic groups, and the total sample. Certain observed differences were tested for significance. In her qualitative analysis she discusses characteristics of the children's responses. In this study similar procedures are used in the analysis of the major sample and the class V subjects.

¹ Probst reported a .94 (corrected to .97) correlation between Forms A and B; the comparable correlation in this study is .89 (corrected to .94).

MILDRED C. TEMPLIN

TABLE I

MEAN REGULAR AND SUBSTITUTE* SCORES ON TOTAL TEST, FORM A
AND FORM B FOR THE MAJOR SAMPLE AND CLASS V
SAMPLES INTERVIEWED IN NOVEMBER AND MAY

	MAJOR-SAMPLE (N = 100)		CLASS V NOVEMBER SAMPLE (N = 24)		CLASS V MAY SAMPLE (N = 24)	
	Mean	SD	Mean	SD	Mean	SD
<i>Regular Score</i>						
Total	58.3	16.2	56.0	14.0	58.7	15.2
Form A	28.5	8.6	27.5	6.4	29.9	8.5
Form B	30.3	8.2	30.6	5.4	29.1	7.1
<i>Substitute Score*</i>						
Total	63.2	17.5	61.2	16.0	63.7	16.7
Form A	31.0	8.7	31.0	8.7	33.6	10.0
Form B	31.6	8.1	30.5	7.7	30.7	7.5

* Score determined by replacing certain Probst questions with new items.

RESULTS

*Quantitative Analysis**Total Scores*

In Table I mean regular and substitute scores on both forms and for the combined forms are presented. The mean total score reported by Probst was 71.6, *SD* 19.05, which is higher at the .01 level of confidence than either regular or substitute total score in this study ($t = 5.25$ for the regular, and 3.24 for the substitute score). This difference cannot be accounted for by poorer test performance during the latter part of the interview in the present study, because the mean scores on Forms A and B do not differ significantly, and Form B actually yields the higher score. Although Probst does not report mean scores for the separate forms, she states that the mean for Form B was 1.6 points higher than the mean for Form A. Nor can the difference between 1928 and 1954 scores be attributed to a difference in the amount of kindergarten experience of the subjects. Although the mean scores on Form A, Form B, and the combined forms tend to be higher in the sample tested in May, in no instance does the difference approach the .05 level of confidence.

The Probst sample yielded higher scores on all categories of knowledge except category X for which no difference occurred, and category VI for which the present sample received the higher score. Using the substitute scores, the differences between the studies are significant at the .01 level for categories I, II, and VIII and at the .05 level for category VII ($t = 5.24$, 3.68, 3.47, and 2.53, respectively).

CHILD DEVELOPMENT

Sex Differences

Table 2 shows that the boys consistently score higher than the girls, but the differences in scores of the two sexes are statistically significant only on three categories for the Probst samples. In both studies boys receive higher scores in six of the same categories, I, III, VI, VII, VIII, and X, but girls in only one, IX. When the scores on the two studies are compared, those of the girls are not significantly different, but those of the boys are significantly higher in the Probst study.

TABLE 2

SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE SCORES OF BOYS AND GIRLS IN THE PROBST AND TEMPLIN STUDIES

	Boys (N = 50)		GIRLS (N = 50)		<i>t</i>
	Mean	SD	Mean	SD	
1. Probst total score	78.0	18.4	67.4	18.7	2.85†
2. Templin regular score	59.6	16.2	56.7	15.9	0.90
3. Templin substitute score* ..	65.6	16.6	60.9	17.3	1.38
<i>t</i> (1 and 3)	3.52†		1.80		

* Score achieved by replacing certain Probst questions with current items.

† Significant at the .01 level.

Socioeconomic Differences

As seen in Table 3, the upper socioeconomic group consistently receives scores which are significantly higher than the lower group. In both studies the higher score of the upper group appears in each category. In the Probst study only the difference for category II is not significant at the .01 level, but in the present study only the difference for category VII reaches this level. The upper Probst sample scores significantly higher than the present upper sample, but the difference between the lower groups is not statistically significant.

Specific Questions²

In comparing the total samples, more subjects in the present study responded correctly on 27 questions, fewer on 98 questions, and the

² A table giving percentages of correct responses on each question by both sexes, both socioeconomic groups, and by the total sample in the present study and by the total sample in the Probst study has been deposited as Document number 5368 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C. A copy may be secured by citing the Document number and by remitting \$1.25 for photoprints, or \$1.25 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

TABLE 3

SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE SCORES OF THE
UPPER AND LOWER SOCIOECONOMIC STATUS GROUPS IN
THE PROBST AND TEMPLIN STUDIES

	UPPER (N = 50)		LOWER (N = 50)		<i>t</i>
	Mean	SD	Mean	SD	
1. Probst total score	80.0	16.5	63.2	17.7	4.90 [†]
2. Templin regular score	64.6	14.6	51.7	14.7	4.41 [†]
3. Templin substitute score*	69.6	15.9	56.9	16.7	3.87 [†]
<i>t</i> (1 and 3)	3.21 [†]		1.83		

* Score achieved by replacing certain Probst questions with current items.

[†] Significant at the .01 level.

same number on nine questions. On 36 questions the differences in the number of correct responses is significant at the .01 level. More Probst subjects responded correctly on 33 and the present-day subjects on three of these. Significant differences appeared on 10 of the questions for which substitute questions were constructed.

New Questions

The new questions are presented in Table 4 with the percentage of correct responses given by both sexes, both socioeconomic groups, the major and class V samples. In several instances, because of the difficulty in constructing comparable questions, more than one substitute question was tried. In determining total scores, the question to which more children responded correctly was used. For example, as a substitute for identifying Jack Dempsey more children identified Rocky Marciano than Joe Louis so the former was used.

Qualitative Analysis

The actual responses of the children are essentially similar to those reported by Probst; yet in both studies the responses reflect the environment in which they were offered. Jackie Coogan was identified as "a boy on Lawrence's sweater," and Dagwood as "a man by our house." The present age is reflected in one child's identification of the part of a plant that is underground as the "shelter." Many identical clang associations are used in both studies. Present-day children still say that butterflies make butter; that bees make beans; that a plumber plumbs or sells plums; etc. Such clang associations still are extended in explanations when children report that a carpenter cleans or lays carpets. Concreteness is also apparent. One example is in responses to "What time is it at noon?": time for lunch, time to go to or return from school, and "When Casey Jones comes on" (a reference to a noon television program).

TABLE 4
PERCENTAGE OF CORRECT RESPONSES ON SUBSTITUTE AND PARALLEL QUESTIONS BY SEX AND SOCIOECONOMIC STATUS

Form	Category	Item Replaced	Question	M A J O R S A M P L E S			C L A S S V			
				Upper (N=50)	Lower (N=50)	Girls (N=50)	Boys (N=50)	Total (N=100)	Nov. (N=24)	May (N=24)
Substitute Questions										
A	III	1	Who is Dagwood?	34	14	28	20	24	17	25
A	III	2	What is a helicopter?	50	44	32	62	47	37	58
A	III	4	Who is Rocky Marciano?	20	20	8	32	20	21	25
A	III	5	Who is Hopalong Cassidy?	94	86	90	90	90	87	100
A	III	6	Who is Dwight D. Eisenhower?	58	44	46	56	51	54	29
B	III	1	Who is Dennis the Menace?	66	56	54	68	61	54	62
B	III	4	Who is George Mikan?	8	6	4	10	14	8	8
B	III	5	Who is Harry Truman?	12	12	8	16	12	8	8
B	III	6	Who is Adlai Stevenson?	10	8	4	14	9	4	4
B	IX	2	What do we have to have to make a cake?	88	78	88	78	83	67	62
B	IX	3	What is a mangle used for?	6	0	4	2	3	0	4
B	IX	4	Where do we keep frozen foods?	98	98	98	98	98	92	96
A	X	6	What makes an iron hot?	14	6	6	14	10	0	12
A	XI	3	You know how to play "The Farmer in the Dell." Whom does the farmer choose?	68	64	80	61	71	46	79
B	XI	1	What do we use to play golf?	72	50	72	50	61	54	50
B	XI	4	What do people play Canasta with?	22	20	24	18	21	29	29
Parallel Questions										
III			Who is Carnival Clown?	68	64	76	56	66	71	71
III			Who is Howdy Doody?	96	96	94	98	96	92	87
III			Who is Cedric Adams?	72	66	64	74	69	79	71
III			Who is Arthur Godfrey?	90	80	88	82	85	83	87
III			Who is Miss Frances?	58	36	52	42	47	46	37

The incorrect responses of children illuminate the levels of their understanding. Lack of appreciation of cause and effect is seen in the response of several children to "How is lumber made from trees?" in which several children started with a board, stuck it into the ground, tied branches on it, put leaves on the branches and ended up with trees. The imaginative explanations of children are illustrated by the idea of one child about lighting: "I think that it's a broken sun."

Only three children did not have television sets in their homes and even they had access to one. The impact of television is apparent in the high percentage of the children who correctly identified both local and national television personalities (see Table 4). It is also evident in the precise, detailed information offered on the program content and schedules, and on the personalities of the performers. "He's on TV," or "He's got a show," was a frequent response in identifying a person a child did not know. The children named their favorite television programs specifically and often described them with a good deal of color and enthusiasm.

In contrast, when asked about their favorite radio programs, the children often named television programs (some of which were also on radio) or general areas such as news, music or sports. Radio was not as universally present or listened to by the children. Where it was, however, it sometimes served a definite function. For example, several children indicated that they listened to "sports with daddy," and one child said "I listen to the weather so my mother knows how to dress me."

That present-day children are, at least verbally, aware of more of the world is illustrated in responses to "What large city is closest to Minneapolis?" The percentage of correct response is about the same in both studies, but only three or four responses reported by Probst mentioned places outside the state of Minnesota—California, London, Chicago, and Washington (which may also refer to a well-known street in Minneapolis). Among the incorrect responses in the present study, two cities outside of Minnesota (Fargo and La Crosse), three states (California, Michigan and Texas), and two countries (Canada and Japan) are mentioned. Although television may influence some responses, in nearly all instances, the youngsters reported some friend or relative living or having lived in the place mentioned. Adults today are more mobile and this is reflected in the responses of the children.

SUMMARY AND DISCUSSION

A quantitative comparison of the responses of kindergarten children in 1928 and 1954-1955 indicates that on the same questions, children in the earlier period scored significantly higher. The differences in the range of information between boys and girls seem to have decreased to a greater extent than those between upper and lower socioeconomic groups.

A qualitative comparison of the actual responses of the children points up essential similarities in child thought between the two periods. In both, children use clang associations, concrete reference, and reflect many aspects of the environment in which they live. There is substantial evidence that television is a pervasive element in the lives of present-day kindergarten children.

Because present-day children gave fewer correct responses to the same questions does not necessarily mean that they are less well informed. It probably indicates that the questions were better suited to the experiences of children in 1928. In constructing the original questionnaire, the greater number of items Probst included dealt "with objects fairly common in the environment of most urban children or with those that form topics of conversation" (3, p. 83). The findings suggest that many such questions may be less pertinent for present-day children. Nevertheless, outside of the category on current topics and history there were few questions which practically none of the children answered correctly. As the total environment changes, the specific environment of the children also shifts, but this environment has changed gradually enough over the past few decades so that a substantial part of the information sampled is not irrelevant.

An important problem inadequately met by both studies is how to describe or analyze the environment in which the study is carried on. That such a problem is inherent in most descriptive studies of child behavior has been recognized (1). An identical question asked children may not mean the same merely because of the time or place it was asked. The question showing one of the greatest shifts in correct responses was "Tell me the name of a Minneapolis newspaper." In 1928 85 per cent gave a correct name, but in 1954-1955 only three children attempted a name, and none knew one. Since most of the children reported newspapers in their homes, the decline is not evidence of unfamiliarity with newspapers. However, in 1928 Minneapolis had two major morning and two major evening newspapers. To identify one, it was necessary to name it. In 1954-1955 the city had one morning and one evening newspaper, so the specific name was less important. In both studies, although the number of "correct" responses differed, it is likely that the children referred to the paper as they heard it referred to about them.

The over-all increased stimulation in the environment may account for the differences in some responses. For example, significantly less information about what was made in the Ford plant was given in 1954-1955. The Ford does not hold the same unique place among automobiles as it did in 1928, soon after the introduction of the new and much-discussed Model A. The St. Paul Ford assembly plant as part of the area's industry does not hold the same unique position that it held when it was three years old and an early example of Ford decentralization in assembling cars. While children of kindergarten age are acquiring much specific information, they are less able to generalize from this information than they will be when

they are a few years older. Fewer children had this particular bit of specific information which was less unique in their environment than it was in 1928.

Change in environmental context was apparent when the substitute questions were being constructed. In the category of current topics and history it was simple to substitute persons in comparable positions for president, etc., although this was no assurance of the comparability of the questions. However, in 1954 there was no popular hero with the appeal of Lindbergh; there was no child actor comparable to Jackie Coogan; and neither Rocky Marciano nor Joe Louis was as well known as Jack Dempsey had been.

The results seem to indicate a greater similarity in the environmental stimulation of present-day children. The decreased differences between the responses of boys and girls are in agreement with other recent findings (4, 5), and may reflect a disappearing double standard in dealing with children. Although still apparent in the present study, significant differences in the performance of children in the upper and lower socioeconomic groups were not as frequent. The number of subjects selected in each occupational class was identical with the Probst sample, although there has been substantial occupational shift and some upgrading since 1928. Had the 1954 occupational distribution of the city been used in selecting the present sample, differences between the performance of socioeconomic groups may have been less apparent. This greater similarity of the environmental stimulation is also suggested by the results of the comparison of the sex and socioeconomic groups in the two studies: insignificant differences were found between the groups obtaining lower scores in the earlier study (girls and lower socioeconomic children), but significant differences were found between the groups obtaining the higher scores on the earlier study (boys and upper socioeconomic children).

The responses of the children in the two studies differ chiefly in the frequency of correct responses. This may merely reflect some shift in the environment of the children. While this series of informational questions gives some indication of how at a given time and place the information of a child "compares with older children his age" (3, p. 83), it cannot necessarily be used to compare children at different places or at different times. To sample adequately the range of information of present-day children an extensive set of current questions must be devised, and, of course, revised continuously or periodically to take into account environmental change.

There is, however, substantial agreement in the qualitative aspects of the responses offered in both studies. The actual responses of the children, frequently the erroneous ones, indicate similarity in performance between the two periods. Had the questions asked in 1928 permitted more comparison of concepts, it may be that this similarity would have extended to more stability in the percentages of acceptable responses as well. The

CHILD DEVELOPMENT

conceptual level attained by children as well as the specific information acquired should be considered in further comparisons over time of knowledge of children. Being able to say the number of pennies in a nickel or a dime is probably at least as much related to specific experiences with money as to having a concept of the quantity "five" or "ten."

REFERENCES

1. BARKER, R. G., & WRIGHT, H. F. *Midwest and its children*. Evanston, Ill.: Row, Peterson, 1954.
2. PROBST, CATHRYN A. Extent and range of information possessed by kindergarten A children in the Minneapolis Public Schools. Unpublished master's thesis, Univer. of Minnesota, 1928.
3. PROBST, CATHRYN A. A general information test for kindergarten children. *Child Develpm.*, 1931, 2, 81-95.
4. TEMPLIN, MILDRED C. *Certain language skills in children: Their development and interrelations*. Minneapolis: Univer. of Minnesota Press, 1957.
5. TEMPLIN, MILDRED C. Play activities of children and adolescents: I. A comparison of urban subjects in 1924, 1950 and 1955. (In preparation)

ADULT REACTIONS TO PREFERRED CHILDHOOD STORIES.¹

MARY J. COLLIER *and* EUGENE L. GAIER
Louisiana State University

Psychologists have tended to neglect the field of children's literature as an area of interest to themselves or, indeed, even to children. Observation, planned or casual, suggests, however, that some stories are so appealing to children that they warrant attention from psychologists on these grounds alone. Particularly striking is a child's intense and persistent preference for his favorite story as if its repetition had important personal meaning to him. One wonders what that meaning is and how one might investigate it.

The psychological literature is of limited help in this connection. Freud (4, 5), Fromm (6), Jones (7), Rank (10), and others have dealt with well-known folk tales and myths and the possibly universal meaning in their recurrent themes and symbols. However, their treatment is chiefly theoretical and without reference to data from people who like the stories, to individual differences in taste, or to circumstances conducive to preference.

Social anthropologists like Benedict (3) and Malinowski (8) deplore the tendency of these writers to interpret myths and folk tales in universal terms. They argue that themes and symbols, apparently similar, take on different meanings in different cultures and that these meanings are never universal. They do claim, however, that regardless of cultural context these stories have the invariable function of codifying and conserving the morals and traditions peculiar to a given society. But, like the psychoanalysts, they indicate little concern for individual reactions.

Accenting literary rather than specifically psychological considerations, Arbuthnot (1, 2) and Meigs *et al.* (9) have discussed the nature and

¹ The writers wish to express their appreciation to Dr. Irwin A. Berg for his enthusiastic support of this project, and to Mr. William D. Leipold for his assistance.

CHILD DEVELOPMENT

function of certain children's stories. They refer also to publishers' and librarians' records as indices of story popularity but remark upon their limited value in establishing the true preferences of children themselves.

The writers are interested in studying the somewhat neglected individual reactions of children to their favorite stories. Therefore, following preliminary work in a nursery school setting, a series of studies was planned for that purpose. Attention was focused on two central questions, "What stories do children like?" and "Why do they seem to like them?" Various approaches, types of subject, and kinds of analysis were planned throughout the series. The present study is the first of the series to be reported.

SUBJECTS AND PROCEDURE

A total of 264 students, 80 men and 184 women, in the Educational Psychology sequence at Louisiana State University were asked to summarize their favorite childhood story, to report the circumstances of first encountering it, and to compare past and present impressions of it.

TABLE I
CLASSIFICATION OF ALL STORIES

Classification	WOMEN		MEN	
	N	%	N	%
Fairy tales	109	59.9	24	30.0
Fiction	49	26.6	32	40.0
Animal stories	16	8.0	15	18.8
Religious stories	7	3.8	8	10.0
Poetry	2	1.1		
Biography	1	.6	1	1.2
Total	184	100.0	80	100.0

RESULTS AND DISCUSSION

The 264 stories obtained were grouped according to Arbuthnot's (1) classification system. They fell into the categories of fairy tales, fiction, animal stories, religious stories, poetry, and biography, in that order of preference (Table 1). Since the overwhelming majority of the subjects (women, 86.4 per cent; men, 70 per cent) chose stories in the first two categories, only these will be discussed here.

The fairy tale category included both the old, anonymously-composed folk tales, fables, myths, and epics, like *Little Red Riding Hood*, *The Grasshopper and the Cricket*, *Jason and the Golden Fleece*, and the so-called

"new tales of magic" like *The Wizard of Oz*, *B'r'er Rabbit*, and *Little Black Sambo*. More than half of the women's preferences were fairy tales with *Cinderella* (chosen by 32) and *Snow White and the Seven Dwarfs* (chosen by 17) clearly the most popular. Less than a third of the men named fairy tales. There was no one fairy story as popular with men as with women. The male selections also showed a contrasting taste for stories in which the chief characters were male or, sometimes, animal, as in *Jack and the Beanstalk* and *The Three Little Pigs*.

Most choices classified as fiction were realistic stories about what Arbuthnot (1) calls the "here and now," that is, "about contemporary life in our own country." Among the men the most frequently mentioned stories were *Tom Sawyer* and *Robinson Crusoe* (seven and five choices, respectively). Most popular among the women were *Heidi*, sometimes called a "modern Cinderella story," *Little Women*, and *The Bobbsey Twins* (eight, four, and four choices, respectively).

The themes, major characters, and story resolutions in the fairy tale and fiction groups will now be considered. Regarding the former, as might be expected, the fairy tale characters, generally, were less realistically portrayed than those in fiction, and the fairy tale resolutions tended to be through magical means. However, there were sex differences in specific choices.

In the fairy tales most often chosen by the women, the heroines like *Cinderella* and *Snow White* tended to be naive, helpless, and persecuted by some malevolent, older woman like a wicked stepmother, witch, or bad fairy. Not through their own actions, but through those of unusual agents like a fairy godmother, a good fairy, or dwarfs, they were protected from their persecutors. Finally, because of their beauty, virtue, or suffering, they were rescued for all time by a prince of somewhat dim personality and with no other apparent function.

The selections of the men, who chose fairy tales about half as often as the women, usually had dominant male or animal characters who were threatened by giants or by impersonal physical dangers, rather than by older women. In fact, when an older woman was present, she was likely to be benign or helpful as in *Jack and the Beanstalk* where the giant's wife protects Jack. Their resolutions were less passive and less magical than those in the women's stories, less, too, a function of virtue, physical attraction, or suffering. Rather, the outcomes were achieved through physical daring, strength, or ingenuity, as in *The Gingerbread Man*, *Pinochio*, *Jason and the Golden Fleece*, *Robinhood*, *The Story of Roland*, and *Hansel and Gretel*.

A consideration of the fiction choices shows that the characters, themes, and resolutions here were generally more realistic than those in the fairy tales. Sex differences, however, persisted.

The most salient feature of the men's stories was the absence of any leading female characters but a pronounced emphasis on one male character

CHILD DEVELOPMENT

and on themes of adventure, of self-assertion, and of overcoming physical obstacles. The chief character in the men's stories tended to be an adult male, alone or in competition with other men, as in stories like *Robinson Crusoe*.

In contrast to the fiction stories chosen by the men, those of the women usually dealt with leading characters who were children, rather than adults; who might be multiple or single; and who might be, unlike the male pattern, either boys or girls. The women's themes dealt chiefly with human relationships, as such, in stories like *Little Women*, *Heidi*, or *Billy and His Aunt Minnerva*; with the mastery of impulses and feelings and with moral values, as in *Oliver Twist* and *The Break in the Dyke*. Themes of adventure, self-assertion or mastery of the environment were infrequent. There was less activity, less solitude, less independence, and less reliance on personal resources in the women's stories but more humor and more emphasis on human limitations, philosophically accepted.

For women, story choice was significantly related to these variables, at the 1 per cent level of confidence: first, that the story was heard at, or before, the age of 6 years; second, that the story had colored pictures; third, that the story had a happy ending (Table 2).

TABLE 2
VARIABLES ASSOCIATED WITH STORY PREFERENCE

Variable	TETRACHORIC COEFFICIENTS	
	Male N=89	Female N=184
Happy ending		
Pictures	.29	.41**
Colored pictures	.26	.24*
Heard at or before the age of 6	.25	.40**
Heard between the ages of 7 and 10	.24	.38**
	.40**	.20

* Significant at the 5 per cent level.

** Significant at the 1 per cent level.

For men, none of the preceding variables was significantly related to story choice; encountering the story between the ages of 7 and 10 years was, however, a highly significant, selective factor.

Several factors might account for this sex difference in the age level at which the preferred story was first heard. First, the women, with earlier language development, might have heard, or read, stories at an earlier age; then, being verbally facile, they might have formulated and remembered them better than men. Second, because of the cultural emphasis on feminine passivity-dependence and on masculine activity-independence,

or because of some biological tendency in this direction, the dependent, childish stage might have been more acceptable and therefore more pleasantly and completely recalled for women. Third, because of a greater cultural emphasis on "realism" for men than for women, the men might have had a greater tendency to forget or conceal any interest in the stories of open make-believe. Fourth, the so-called "Oedipal" and "latency" periods usually correspond to the 3 to 6 year and the 7 to 10 year spans, respectively. These were preferred story periods for women and men, respectively, suggesting that the women might have been more accepting of, or were still more involved in, conflicts stemming from the Oedipal period and the men, those of latency.

That the women should prefer stories with happy endings more frequently than the men might be related to these factors. First, they chose stories from the younger age period where stories may simply have happy endings more frequently than do those for older children. Then, also, stories with happy endings were perhaps read, or given, to girls more often than to boys for "protective" reasons. Perhaps, also, regardless of their story repertoire, girls may have simply preferred happy fantasy more than boys either because it was a more congenial, or a more approved, method of mastering their particular anxieties. The more active methods were possibly more congenial to, or culturally more approved for, the boys.

That women should prefer stories with illustrations, especially those that are colored, might be explained by the fact that stories for younger children, most often chosen by the women, simply are more copiously and colorfully illustrated than those for the later age groups. Another possibility is that cultural patterns may differentially reinforce men and women for attention to decorative detail, or a native sex difference of this kind may be operating. Perhaps, too, the preferred story, through "halo effect," may be endowed retrospectively with colorful illustrations by women more often than by men because of a differential flair for, or value placed upon, them.

Negative correlations (significant at less than the 1 per cent level) were found between age at which the story was first encountered and the age for which it was recommended for both men and women (men, $-.82$; women, $-.81$; total group, $-.83$). Thus, the preferred story was usually recommended as appropriate for a child older than the subject when he first encountered it. Speculations about this relationship follow. Perhaps the subjects were, or saw themselves as, precocious; hence, they adopted the attitude, "Other children could not appreciate my favorite story as early as I did." It is also possible that they may have made some consistent initial error in their age-placement of the story and then made a consistent effort to revise the error in prescribing an appropriate age. Perhaps, too, as children, the subjects might have selected a favorite story which initially was more appropriate for an older child and then made a corresponding allowance in retrospect. Finally, the story may have been disturbing,

CHILD DEVELOPMENT

though preferred, its disturbing quality later attributed to a premature encounter with the story and the recommendation then made that it be used for older children.

When present and past attitudes toward the preferred story were compared, they appeared generally consistent with one another since about two-thirds of both men and women considered their old favorite to be their present preferred story for children (Table 3). A few had changed their minds and the remainder were uncertain about their preference for the story now. Possible explanations for this finding are as follows: First, past and present attitudes might have been consistent, as reported, or the report might have reflected, rather, some effort to maintain the appearance of consistency. Second, although the stories were preferred in the present, as in the past, adult reasons for liking them might have differed from those of childhood. Third, previous and past judgments may simply have contaminated one another.

TABLE 3
RESPONSES TO THE QUESTION: IS THIS STILL THE CHILDHOOD
STORY YOU LIKE BEST?

Response	M A L E		F E M A L E		T O T A L	
	N	%	N	%	N	%
Yes	55	68.8	121	65.8	176	66.7
No	3	3.7	10	5.4	13	4.9
Uncertain	22	27.5	53	28.8	75	28.4
Totals	80	100.0	184	100.0	264	100.0

The possibility was indirectly explored that other people who liked the story exerted an influence on these subjects. However, about one-third of the women and almost one-half of the men indicated that the stories they had preferred as children had not, then, been the favorites of anyone else. Even so, almost half of both sex groups reported that their preferred stories had also been the favorites of siblings and peers. Only a few subjects mentioned, as significant, the attitudes entertained by older persons toward their favorite stories; when they did so, the older persons were almost invariably women. The results have certain implications. Men were, or recalled that they were, more independent of others' attitudes in selecting their favorite story in childhood. The preferences of siblings and peers carried more apparent weight in determining our subjects' choices than did those of older persons. Finally, when the judgments of older persons did affect the subjects' childhood choices, they were likely to be those of an older woman rather than of an older man.

SUMMARY AND CONCLUSIONS

In summary, this study is one of a series planned to explore the responses of children to their favorite story at several stages of development. Its purpose was to study the characteristics of stories preferred in childhood by college students and evaluated in retrospect.

The subjects, 264 college students, were asked to recall and summarize their favorite childhood stories and to answer questions about past and present impressions of them. Responses were analyzed for differences in age at which the story was encountered, major characters, persons introducing the stories, major themes, and differential treatment of similar themes.

Results show that women chose stories stemming from the Oedipal period, involving evil mother figures, benign but active males, and persecuted, passive young women. Their stories had been shared more with loved persons, covered a restricted number of themes, and had happy, often magical, endings. Men preferred stories encountered independently during the latency period with themes of adventure, problem solution, and self-assertion. Story endings and evaluations of the stories were more obviously reality-oriented in the male group; women figures, though rare, were almost exclusively kind and maternal.

In conclusion, types of story preferred and the manner of reporting them appear to reflect cultural, and possibly biological, sex-role expectancies.

REFERENCES

1. ARBUTHNOT, MAY H. *Children and books*. New York: Scott, Foresman, 1947.
2. ARBUTHNOT, MAY H. *Time for fairy tales*. New York: Scott, Foresman, 1952.
3. BENEDICT, RUTH. FOLKLORE. In *Encyclopedia of the social sciences*. New York: Macmillan, 1937. Vol. 5.
4. FREUD, S. *Introduction to psychoanalysis*. New York: Boni & Liv-
eright, 1920.
5. FREUD, S. *The interpretation of dreams*. London: Hogarth, 1953.
6. FROMM, E. *The forgotten language*. New York: Rinehart, 1951.
7. JONES, E. Psychoanalysis and folklore. In E. Jones (Ed.), *Essays in
applied psychoanalysis*. Vol. II. London: Hogarth, 1951.
8. MALINOWSKI, B. *Myth in primitive psychology*. New York: W. W.
Norton, 1926.
9. MEIGS, C., EATON, A., NESBITT, E., & VIGUERS, R. *A critical history
of children's literature*. New York: Macmillan, 1953.
10. RANK, O. The myth of the birth of the hero. *J. nerv. ment. Dis.*, 1914.
(Monogr. Series No. 18)

A TEST ANXIETY SCALE FOR CHILDREN¹

SEYMOUR B. SARASON, KENNETH DAVIDSON, FREDERICK
LIGHTHALL *and* RICHARD WAITE

Yale University

For a number of years studies were done at Yale by Sarason and others (1, 2, 4, 5, 6, 7, 10, 11, 12, 13) on the effects and correlates of test anxiety. In those studies subjects were college students. For a variety of reasons it was decided to study this anxiety in school age children. Among the major reasons were (a) the unduly selected nature of a college population, limiting whatever generalizations one can make, (b) the practical difficulties one faces in securing the continual cooperation of such subjects, (c) the opinion that test anxiety was a reaction which long antedated entrance to college, (d) indications in the literature (8; 9, p. 168) that fear of school failure was one of the most common of worries or fears among children, and (e) the hope that the valid discernment of this anxiety would have more than a little practical and clinical value in the school setting. The present report concerns the initial findings of a long term project on the measurement and correlates of anxiety in children of school age. It is hoped that other research workers may find the test anxiety scale a helpful tool.

PROCEDURE

The initial form of the childrens' test anxiety scale consisted of 43 items concerned with attitudes toward and experiences in test and test-like situations.² As in the case of the college scale, but never made explicit in

¹ The project of which this study was a part is being supported by a grant from U.S.P.H.S. We wish to thank Mr. Joseph Foran, Superintendent of Schools in Milford, and Dr. Elmer Hagman, Director of Research and Guidance in Greenwich, for the support which they gave to the project. We owe a particular debt of gratitude to the teachers and principals of the schools used for the time and effort they gave in the collection of data. The teachers contributed a good deal of their own time in making ratings, for which we are indeed grateful. We wish also to thank Dr. John Doris for his assistance. At all stages of this project Dr. Robert Abelson was of great help to us in statistical matters.

² Copies of the shortened form of the scale for research purposes are available upon request. This revised form consists of 30 of the items of the original scale. In subsequent studies with the shortened scale the same results as those reported in the present article were found. We chose to present the data on the longer scale because a much larger number of cases were used and give one a better picture of the stability of the findings.

CHILD DEVELOPMENT

the published articles, the choice of items was guided by consideration of Freud's (3) three criteria of the anxious reaction: it is unpleasant, there are physiological concomitants, and it involves conscious awareness—*anxiety* is a conscious experience. Below are several of the items:

When the teacher says she is going to find out how much you have learned, does your heart begin to beat faster? (Hand on heart—illustrate).

When the teacher wants to find out how much you have learned, do you get a funny feeling in your stomach? (Hand on stomach—illustrate).

When the teacher asks you to write on the blackboard in front of the class, do you sometimes feel that the hand you write with shakes a little?

Do you worry a lot *before* you take a test?

Do you worry a lot *while* you are taking a test?

After you have taken a test, do you worry because you think you did not do well?

When the teacher says she is going to give the class a test, do you become afraid that you are going to fail—that you will do poorly?

When the teacher says she is going to give the class a test, do you get a nervous (or funny) feeling?

While you are taking a test, do you usually think you are not doing well?

The scale was group administered. Each item was read to the class and the children were instructed to encircle a "yes" or a "no" on answer sheets. Preceding administration of the scale the following was said: "We are not giving you a test. When the teacher gives you a test, there are right and wrong answers. We are going to ask you questions about how you feel and what you think but there are no right or wrong answers. When we ask you a question, think about it and then put a circle either around yes or no. Maybe you will put a circle around yes and your neighbor puts one around no—that is because you feel differently. It does not make any difference what anybody else does. You think about each question and then put a circle around a yes or a no. Remember, there are no right or wrong answers. We want to know what you think and feel. Try to be as honest as possible." The scales were always administered by non-school personnel. The teacher of the class was never present. The children were told that no one but the person administering the scale would ever see the child's answer sheet. In addition to the test anxiety (TA) scale each child was rated by his teacher on a 17-item scale of anxiety behavior in the classroom. Below are some of the items which were given to the teachers:

Does the child exhibit unwarranted fidgeting (e.g., squirming, restless behavior) when called upon to recite in class?

Does the child tend to become upset or anxious when a test is announced in class or he is called upon to recite?

Does the child function better when working alone than when working before the class?

Does the child work better in a situation in which he can take his time than in one in which there is time pressure?

Finally, each child was given the Otis Alpha group intelligence test as well as the Stanford Achievement test.

Before proceeding further, it seems appropriate to describe briefly certain reactions of the children to the TA scale. First, when the examiner walked into the classroom carrying the answer sheets, it almost always happened that at least one child said: "We're going to get a test!" This was even true in the earliest grades where the advisability of giving the scale had been questioned because tests were never given. Equally as striking was the fact that following administration of the TA scale, some children could be counted on to ask: "When will we find out how we did?"—this despite the instructions. These incidental observations indicate how early in life children in our culture become test-oriented and suggest how pervasive and interfering such an orientation may be. By interfering, we here mean the tendency to adopt a set which is explicitly not called for in the situation.

TABLE I
BREAKDOWN OF TOTAL SAMPLE BY GRADE AND SEX CATEGORIES

Sex	Grade 2	Grade 3	Grade 4	Grade 5	Totals
<i>Milford</i>					928
Boys	277	205	241	205	769
Girls	221	181	193	174	1697
Total	498	386	434	379	

SUBJECTS

Children in grades 2 through 5 in six schools in nearby Milford were studied. Table 1 breaks down the subjects by grade and sex. Two schools in Greenwich ($N = 600$) were also used but, since the results are essentially the same for the two school systems, only the results for Milford will be presented. It is important to emphasize that although Milford and Greenwich differ markedly in a number of ways (e.g., social class of the population), the findings in both systems were essentially the same.

RESULTS

Reliability. Table 2 presents the findings for test-retest (two-month interval) reliability. Table 3 presents the data for split-half reliability. For both procedures a satisfactory degree of reliability is obtained.

CHILD DEVELOPMENT

TABLE 2

TEST ANXIETY QUESTIONNAIRE TEST-RETEST RELIABILITY

Grade	No. of Classes	No. of Subjects	r	SE_r^*
5	3	73	.82	.040
4	3	77	.76	.049
3	2	54	.44 [†]	.110
2	4	90	.76	.045
All grades	12	294	.71	.029

* SE_r (after Soper in Snedecor).

[†] One S obtained a score of 1 in test and a score of 41 on retest. If this S is omitted r for third grade becomes .65 and $SE_r = .08$.

TABLE 3

TEST ANXIETY QUESTIONNAIRE SPLIT-HALF RELIABILITY

Grade	No. of Classes	No. of Subjects	r (uncorrected)	r^* (corrected)
5	8	208	.788	.881
4	8	215	.795	.886
3	8	204	.816	.899
2	8	198	.695	.820

* Corrected by Spearman-Brown formula (Peatman).

Increase in TA score with grade. It was anticipated that TA scores would tend to increase with grade. This tendency was found to be statistically significant (Table 4) although it is clear that the increase in score is not a regular or steady one.

Correlation between TA and IQ. On the basis of previous studies with college students, as well as the assumption that the test anxious reaction

TABLE 4

CHI SQUARE COMPARISON OF TEST ANXIETY SCORES BY GRADE
(Milford School System, $N = 1697$, Mdn. Score = 12-13)

	G R A D E				χ^2	p
	2	3	4	5		
No. of cases above median	213	192	198	226	26.9	< .01
No. of cases below median	285	194	236	153		

was primarily interfering in its effect, it was anticipated that the correlation between TA and IQ would be small, negative, but significant.³ Going from grade 2 to 5 the correlations, all significant at the 1 per cent level, are $-.196$, $-.214$, $-.275$, and $-.284$. These findings are essentially the same as those reported by Zweibelson (14) with the Otis Alpha and Beta. He found, in addition, that the TA correlated significantly less with the Davis-Eells Problem-Solving games than with either the Otis Alpha or Beta—a finding predictable from the rationale of the Davis-Eells Games. Because of the high correlation between IQ and mean achievement score in the present sample, it is not surprising that the correlations between TA and mean achievement are very similar to those reported above. Going from grades 2 to 5 the correlations are: $-.186$, $-.249$, $-.270$, and $-.294$ (all significant at the 1 per cent level).

TABLE 5

CORRELATIONS OF TEST ANXIETY AND TEACHERS' RATINGS, CLASS BY CLASS, AVERAGED FOR THE VARIOUS GRADES IN BOTH SCHOOL SYSTEMS

Grade	No. of Classes	Total N	Average r^*	Significance of r p
<i>Greenwich</i>				
2	6	142	.31	< .001
3	7	160	.21	< .02
4	6	142	.29	< .001
5	6	139	.27	< .01
All grades	25	583	.27	< .001
<i>Milford</i>				
2	17	469	.10	< .05
3	15	377	.09	< .10
4	15	417	.34	< .001
5	14	365	.18	< .01
All grades	61	1628	.18	< .001
<i>Total Both Schools</i>				
All grades	86	2211	.20	< .001

* The average r was obtained by Fisher's method of transforming the original r for each class into z and then obtaining the weighted average value of z and reconvert to r .

Correlation of TA with teachers' ratings of anxious behavior. Although we realized the difficulties involved in utilizing teachers' ratings, it was felt that it would speak for the face validity of the TA scale if a significant rela-

³ In the college study (11) where this finding was reported the correlation between TA and IQ was reported as $+.209$. A study of the scattergram given in that article will indicate that the positive correlation is a function of how the scattergram was constructed. It is clear from the scattergram that high TA goes with low IQ score.

CHILD DEVELOPMENT

tionship could be found between TA (a child's self-rating, so to speak) and his teacher's independent rating. Table 5 presents the data indicating that a significant relationship does obtain. It will be recalled that in the previous paragraph it was reported that the correlations between TA and IQ (and between TA and mean achievement) tended to increase as one went up the grades. The opposite occurs in the case of the correlations between teachers' ratings and IQ.

	Grade 2	Grade 3	Grade 4	Grade 5
TR vs. mean achievement	-.419	-.353	-.361	-.127
TR vs. IQ	-.301	-.300	-.176	-.125
TA vs. mean achievement	-.186	-.249	-.270	-.294
TA vs. IQ	-.196	-.214	-.275	-.284

In Greenwich the decrease with grade in the correlation between teachers' ratings and mean achievement is even more striking, although a similar steady decrease in the correlation with IQ is not found. It would appear, therefore, that, whereas the teachers' estimate of a child's anxious reaction decreasingly predicts a performance criterion, the children's self-estimate tends increasingly to predict the criterion. It may be that in making their ratings teachers are largely influenced by how they think children are actually performing academically and that such judgments are far more valid in the early, in contrast to the later, grades. It may also be that there are more (and more glaring) overt indicators of anxiety in the earlier grades, i.e., the anxious child in grade 2 is more easy to spot by the teachers than the anxious child in grade 5.

In the course of subsequent studies, as well as on the basis of discussions with many teachers, we have come to the opinion that teachers frequently find it difficult to believe that a bright child may also be one who sees himself as anxious about performance. If a child is considered by the teacher as an average or below average performer, the teacher seems to be much more aware of the possibility of the presence of anxiety than in the case of the clearly bright child whose classroom problem-solving behavior appears perfectly adequate to the observer. If there is in fact less chance for the anxiety of the bright child to be recognized, there would seem to be less chance for it to be reduced in strength.

TA and occupational level. The occupational level of the parent was obtained from information given by the parent when the child was first entered in school. In each case where such information was available, the parent's occupation was put in one of seven categories—1 representing the highest and 7 the lowest category. There was satisfactory agreement between two independent raters in making these categorizations. However, it was clear that the few words (or the one word) given by the parent in describing his occupation was frequently unrevealing insofar as type or

complexity of job was concerned. The average correlation (r transformation) between TA and occupational level is .119 which is significant at the .001 level. It seems reasonable to conclude that occupational level is *one* of the factors significantly related to TA. That this correlation is not larger may well be due to the inadequacies of the raw data mentioned above.

The relationship of TA with reading, arithmetic, and behavior problems. The "diagnosis" of an academic and behavior problem was made by the teacher. A chi square analysis (problem vs. no-problem cases) revealed that individuals who had either a reading, or an arithmetic, or a behavior problem obtained significantly higher TA scores than no-problem cases.

DISCUSSION

The expected increase in TA score with grade, the negative correlation of TA with IQ, the significant correlation between TA and teachers' ratings—these and other results given above support the conclusion that test anxiety is a measurable and meaningful variable. When one considers the pervasive use of tests in our culture and the ways in which they determine the lives of the people who take them, it would be expected that the testing situation would engender an anxious reaction in more than a few individuals.

The significant and negative correlation between TA and IQ is deserving of special discussion. At first glance this finding might be interpreted simply as indicating that the less intelligent (i.e., lower test scoring) individual experiences anxiety in testing and test-like situations because it is and has been a situation in which he encounters difficulty and failure. There are several factors which indicate that such an explanation may be an oversimplification. For one thing the negative correlation was also obtained in the two Greenwich schools which were explicitly chosen because they represent an above average group intellectually. While the above explanation could still be offered for the Greenwich sample, it is clear that the term "less intelligent" could not be used in any absolute sense but in a relative way, i.e., relative to the group studied. In the previously studied college groups, which were clearly intellectually superior and where there was also a restricted range of test scores, the negative correlation also was found. In addition, whereas the high anxious college students received significantly lower aptitude test scores than the low anxious students, the former received significantly higher course grades than the latter. Zweibelson's study with a Milford sample rather strongly suggests that the size of the correlation between TA and IQ varies as a function of the type of test, mode of administration, etc.—findings in keeping with those obtained with the college groups.

Still another consideration which cautions one against simple explanations of the negative correlation between TA and IQ lies in the entirely

reasonable assumption that the IQ score, obtained as it is in a test situation, already reflects the interfering effects of anxiety. It is precisely because of this assumption that any explanation of the negative correlation between TA and IQ must at least consider the difference between such terms as "the less intelligent" and the "lower test scoring" individual.

There are, of course, many bright children (and college students) who have high TA scores the sources of which are not explained by any intellectual inadequacy. It is one of the main points of interest of this project to attempt to understand the sources of high anxiety in children who, from the standpoint of objective indices of performance and achievement, have experienced "success" rather than "failure."

We have here presented initial findings on a test anxiety scale for children. This scale was developed as part of a long term project concerned with the development of anxiety in children. Other studies under way concern the relationship of test anxiety to learning under varied conditions, to direct observations in the classroom, to parental attitudes and report, to behavior in the Rorschach situation, to other anxieties, and to the assignment of blame in failure situations. In addition cross-cultural comparisons are being carried out in England and Australia.

REFERENCES

1. COX, F. N., & SARASON, S. B. Test anxiety and Rorschach performance. *J. abnorm. soc. Psychol.*, 1954, 49, 371-377.
2. DORIS, J., & SARASON, S. B. Test anxiety and blame assignment in a failure situation. *J. abnorm. soc. Psychol.*, 1955, 50, 335-338.
3. FREUD, S. *The problem of anxiety*. New York: W. W. Norton, 1936.
4. GORDON, E. M. Self esteem, need for achievement, test anxiety and performance. Unpublished doctor's dissertation, Yale Univer., 1955.
5. GORDON, E. M., & SARASON, S. B. The relationship between "test anxiety" and "other anxieties." *J. Pers.*, 1955, 23, 317-323.
6. MANDLER, G., & SARASON, S. B. A study of anxiety and learning. *J. abnorm. soc. Psychol.*, 1952, 47, 166-173.
7. MANDLER, G., & SARASON, S. B. The effect of prior experience and subjective failure on the evocation of test anxiety. *J. Pers.*, 1953, 21, 336-341.
8. PINTNER, R., & LEV, J. Worries of school children. *J. genet. Psychol.*, 1940, 56, 67-76.
9. PRESSEY, S. L., & ROBINSON, F. P. *Psychology and the new education*. New York: Harper, 1944.
10. SARASON, S. B., & GORDON, E. M. The test anxiety questionnaire: scoring norms. *J. abnorm. soc. Psychol.*, 1953, 48, 447-448.

11. SARASON, S. B., & MANDLER, G. Some correlates of test anxiety. *J. abnorm. soc. Psychol.*, 1952, 47, 810-817.
12. SARASON, S. B., MANDLER, G., & CRAIGHILL, P. G. The effect of differential instructions on anxiety and learning. *J. abnorm. soc. Psychol.*, 1952, 47, 561-565.
13. WIENER, G. The interaction among anxiety level, stress instructions and level of difficulty. Unpublished doctor's dissertation, Yale Univer., 1954.
14. ZWEIBELSON, I. Test anxiety and intelligence test performance. *J. consult. Psychol.*, 1956, 20, 479-481.

EXTINCTION OF THE SMILING RESPONSE IN INFANTS AS A FUNCTION OF REINFORCEMENT SCHEDULE¹

YVONNE BRACKBILL
*Stanford University*²

Scheduling of reinforcement has been investigated extensively in relation to the conditioning of non-social responses. The results of these studies indicate in general that a schedule by which reinforcement follows every response is less effective in maintaining performance of that response during extinction than is a schedule by which reinforcement follows only some of those responses (4).

This investigation attempts to extend the study of frequency and patterning of reinforcement to the area of social learning and to cortically immature subjects. The purpose of the research is to evaluate the relative efficacy of intermittent as opposed to regular reinforcement upon frequency of smiling in infants.

The use of an instrumental conditioning paradigm in this study is in contrast to past studies of the conditioning of the smiling response (2, 7, 8), all of which, though differing in methodology, have utilized a classical conditioning paradigm in conceptualizing results. The inappropriateness of this framework in accounting for response change over time is most clearly illustrated in the study by Dennis (2).

METHOD

Subjects

Eight normal infants between the ages of three and one-half to four and one-half months served as Ss. Six were males and two, females. All came from intact middle class homes located within 10 miles of Stanford University.³

¹ This report is based upon a dissertation submitted to the Department of Psychology, Stanford University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer wishes to thank Professors Albert Bandura, Quinn McNemar, Frances Orr Nitzberg, Robert Sears, William Verplanck, and C. L. Winder for their suggestions and help.

² Now at National Institute of Mental Health, Bethesda, Maryland.

³ The author is greatly indebted to Drs. John Anderson and Bruce Jessup for their cooperation in referring subjects.

This particular age range was selected with the consideration in mind that the subject had to be old enough to remain awake for a short time period after feeding, yet not old enough to respond differentially to "mother" vs. "others." Other requirements for selection of Ss were: (a) that the infant not cry so often, with such intensity, and for so long a time that sessions very frequently had to be terminated, with consequent lack of experimental progress; (b) that the infant show an operant rate of at least two responses per five-minute interval; and (c) that the infant be able to maintain a supine position for intervals of five minutes without persistent struggling to regain the prone position.

Experimental Procedure

Ss were assigned to the regularly reinforced (RR) group or to the intermittently reinforced (IR) group in consecutive order of their acceptance as subjects. The first S was assigned to the IR group; the second, to the RR group; and so on.

For both groups, the experimental procedure was divided into three periods: In the first or *operant* period E stood motionless and expressionless at a distance of approximately 15 inches above S, and observed him for eight five-minute intervals to ascertain the operant level of smiling. In the second or *conditioning* period, reinforcement was meted out, contingent upon S's smile. Specifically, as soon as S smiled,⁴ E smiled in return, began to speak softly to S, and picked it up. S was then held, jostled, patted, and talked to for 30 seconds before being replaced in the crib.⁵ In the third or *extinction* period, the procedure was again observation without reinforcement, as in the operant period. During extinction, Ss were observed for 13 five-minute intervals.

During all three periods, the basic interval for determining rate of responding was five minutes, and in the discussion to follow, *five-minute interval* (or simply, *interval*) refers to one five-minute period of continuous experimentation. The term *rate of response* refers to number of responses per interval. Also, the terms *response* and *trial* will be used interchangeably.

The two groups of Ss differed in respect to the reinforcement schedules used within the conditioning period. The RR group was maintained on a regular reinforcement schedule during the entire conditioning period. The IR group was maintained on a regular reinforcement schedule until each S had responded at maximum rate (*see below*) for 10 consecutive intervals. Immediately after the tenth interval at maximum rate, Ss were switched to a 2:1 randomized variable ratio reinforcement schedule for

⁴ Interjudge reliability concerning the decision as to whether a smile had or had not occurred was found to be .975, using a total N of 970 such decisions in the formula: twice number of agreements / total number of judgments. This reliability check was done prior to the experiment, with the assistance of Mr. Thomas Milburn.

⁵ The reinforcement procedure was recorded on 12 feet of 8 mm. film.

a total of 60 responses (hence, 30 reinforcements), then to a 3:1 schedule for a total of 45 responses, and finally to a 4:1 schedule for 20 responses. Separate randomizations were used for each subject.

Maximum rate of response was determined as follows. Forty-five seconds were required for the administration of each reinforcement plus its accompanying events. (Five seconds were required for picking *S* up; 30 seconds for reinforcement; five seconds for putting *S* down; and five seconds for recording.) Therefore, no more than six responses could occur and be reinforced during any five-minute interval. *Maximum rate of response* was defined as no fewer than four responses per interval. That particular trial or response that marked the beginning of maximum response rate will be referred to as *criterion*.

Because of intersubject variation in total number of responses to criterion and because it was desired that the experimental results be a function only of reinforcement schedule and not a function of reinforcement schedule plus number of emitted responses, it was necessary to match the two groups for total number of trials during the conditioning period. The method of matching—like the method of group assignment mentioned earlier—was an individual one and was done in consecutive order of admission to the experiment. Specifically, *S* No. 2 (RR group) was matched for total number of trials with *S* No. 1 (IR group); *S* No. 4 was matched with *S* No. 3, and so on.

General Procedure

Each infant was placed on a schedule of social deprivation during the entire experimental period. Social deprivation was defined as minimized social and body contact between the infant and any other person in its environment, excluding *E*. In effect, this meant elimination of all social and body contacts that were not absolutely necessary for *S*'s well being.

Experimental conditions were kept as standard as possible for all sessions and for all *Ss*. *E* always wore a white laboratory coat. *S* was placed in a supine position near the open edge of his crib; placement of *S* and the crib was the same for all sessions. Source of light was standardized, and variance in intensity minimized. At the beginning of each session, *S* was freshly diapered, had just eaten to satiety, and had been awake for from 15 to 20 minutes. During the operant and extinction periods, at the end of every five-minute interval, *S* was placed in a prone position to rest for three minutes. While he was being turned over and during the rest period, *E* was not in his field of vision. During the experimental period, *S* was frequently required to work without rest for two successive intervals, but not for longer.

Typically, there were two to three sessions per day. Length of any one session ranged between 10 and 60 minutes, and was a function of the length of time *S* remained awake. The total number of days spent in experimental

CHILD DEVELOPMENT

tion with any one *S* ranged from eight to 16. Because of the very large amount of time per day required for travel, experimentation, and maintenance of rapport with the mother, only one *S* was run at a time.

RESULTS AND DISCUSSION

Performance During the Operant Period

The two groups did not differ significantly in either mean operant response rate ($t = .29, 7 df$) or total number of responses emitted ($t = .06, 3 df$). For this reason, scores have been combined for graphic presentation (Fig. 1).

Figure 1 also includes a cumulative plot for a ninth, control *S*, who was run, without reinforcement, for an extended operant period of 19 intervals—or roughly three times the length of the operant period for experimental *Ss*. These data provide some evidence against the possible explanation that the mere presence of *E*—and not the reinforcement—is conducive to increased rate of response.

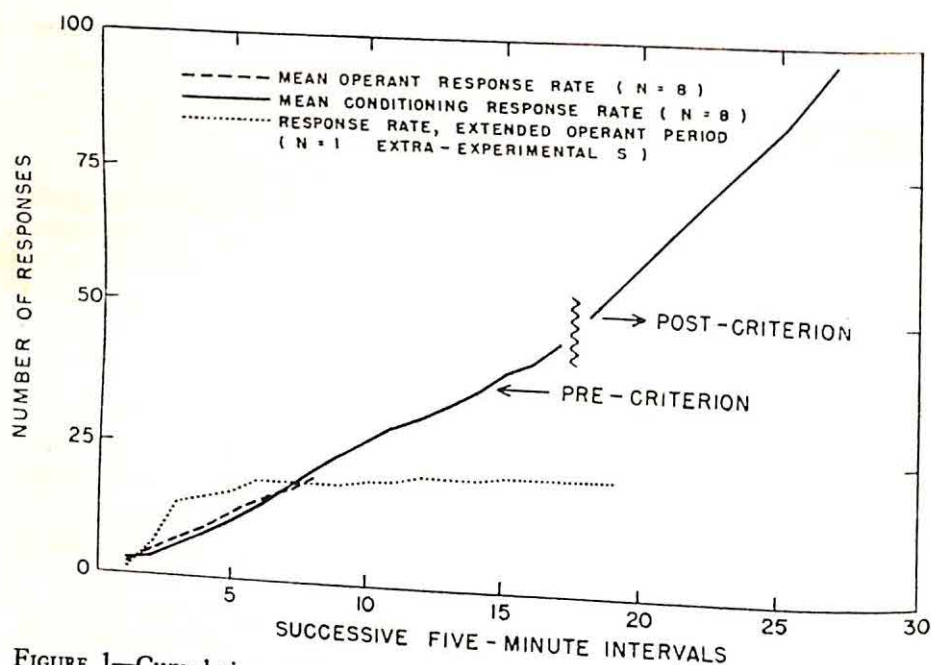


FIGURE 1—Cumulative curves showing rate of smiling response during operant and conditioning periods.

Performance During the Conditioning Period

For the conditioning period, the total number of responses for each group was 927, while the mean number of responses per subject was 231.7. Mean group response rates are presented in Table 1; the most noticeable

features of these data are the abruptness of the change in mean response rate and in variability following criterion.

Relative stability in response rate is also reflected in Table 2, which contains the individual post-criterion response rates emitted under the regular reinforcement schedule. For the IR Ss, percentages are based on the 10 intervals of regular reinforcement following criterion. For the RR Ss, percentages are based on the total number of intervals following criterion. From this table it can be seen that although most Ss habitually worked at one rate of response, there were interindividual differences in the habitual or "preferred" rate.⁶

TABLE 1

GROUP RESPONSE RATES DURING OPERANT AND CONDITIONING PERIODS

	CONDITIONING RESPONSE RATE							
	OPERANT RESPONSE RATE		Preceding Criterion		For First 10 Intervals Follow- ing Criterion		For 11th-22nd Intervals Follow- ing Criterion	
	Mean	σ	Mean	σ	Mean	σ	Mean	σ
Regular Group	2.11	.63	2.70	.47	5.15	.06	5.15	.06
Intermittent Group	2.88	.36	2.43	.44	5.10	.06	6.32* 1.11* 8.12† .64† 13.00‡ 1.51‡	
Combined Groups	2.49	.64	2.56	.50	5.12	.06		

* 2:1 schedule.

† 3:1 schedule.

‡ 4:1 schedule.

Performance During the Extinction Period

Two statistical tests were applied to the extinction data in order to test the hypothesis that the IR group would be more resistant to extinction. First, McNemar's pseudo three-way analysis of variance (5), with blocks representing the experimental variable of reinforcement schedule, was applied to the response frequencies over all 13 extinction intervals. (Summing

⁶ One last datum concerning performance during conditioning was provided by S No. 8, who gave a clear demonstration of the type of response that Skinner has termed "superstitious behavior" (6). The behavioral sequence was as follows. During reinforcement, S kept his left fist doubled in his mouth. When placed in the crib, he withdrew the fist from the mouth and kept it suspended in air for the short time it took him to smile. Then, simultaneously with the beginning of the reinforcement procedure, the fist was promptly reinserted into the mouth, the head turned 90 degrees to the left, and the body musculature stiffened. The onset of this stereotyped response coincided with criterion (56th trial). It disappeared for three days during a period of illness (90th through 118th trials), reappeared in full strength with recovery (119th trial), gradually diminished in intensity, and finally disappeared altogether by the 162nd trial.

CHILD DEVELOPMENT

TABLE 2

INDIVIDUAL RESPONSE RATES DURING CONDITIONING: PERCENTAGE OF INTERVALS FOLLOWING CRITERION DURING WHICH *S* EMITTED FOUR, FIVE, OR SIX RESPONSES PER INTERVAL

		Percentage of These Intervals During Which Response Rate Was		
Total Number of Intervals on Regular Reinforcement Following Criterion		4	5	6
<i>Intermittent Group</i>				
<i>S</i> No. 1	10	10%	80%	10%
<i>S</i> No. 3	24	20	70	10
<i>S</i> No. 5	10	30	10	60
<i>S</i> No. 7	45	20	40	40
<i>Regular Group</i>				
<i>S</i> No. 2	10	4.2	58.3	37.5
<i>S</i> No. 4	31	13.3	62.2	24.4
<i>S</i> No. 6	10	80.6	19.4	0
<i>S</i> No. 8	22	0	22.7	77.3

across 13 intervals, the total number of responses for the IR group was 331; for the RR group, 130.) The resulting F of 17.14, with 1 and 6 df , is significant beyond the .01 level. Second, a mean difference was computed, using the four matched-pair differences in total number of responses during the last six extinction intervals. (In this case, total number of responses for the IR and RR groups were 113 and 3, respectively.) For 3 df the t of 6.77 is significant beyond the .005 level by a one-tailed test.

A noncumulative plot of mean response rates during extinction (Fig. 2), shows some interesting periods of sharp rise in response rate for both groups (interval No. 5 for the RR group; intervals No. 7 and No. 8 for the IR group). The point of occurrence of this "recovery" did not appear to be a function of the length of preceding rest period; e.g., in the IR group, two *S*s recovered after a rest of several hours (i.e., at the beginning of a new session), while the other two *S*s recovered after a rest of only three minutes (i.e., during a session).

Figure 2 also indicates that every member of the RR group extinguished not to his previous, operant rate of response, but to a zero rate. Coincident with the beginning of zero response rate was a conspicuous behavioral change: *S* would no longer fixate the discriminative stimulus (*E*'s face). Instead, *S*'s head turned to one side and remained there—an occurrence, it might be pointed out, that was in distinct contrast to *S*'s persistent fixation during conditioning. When this occurred, immediately preceding the last extinction interval, *E* propped *S*'s head with rolled blankets or other material, making it impossible for the infant to turn his head to the left

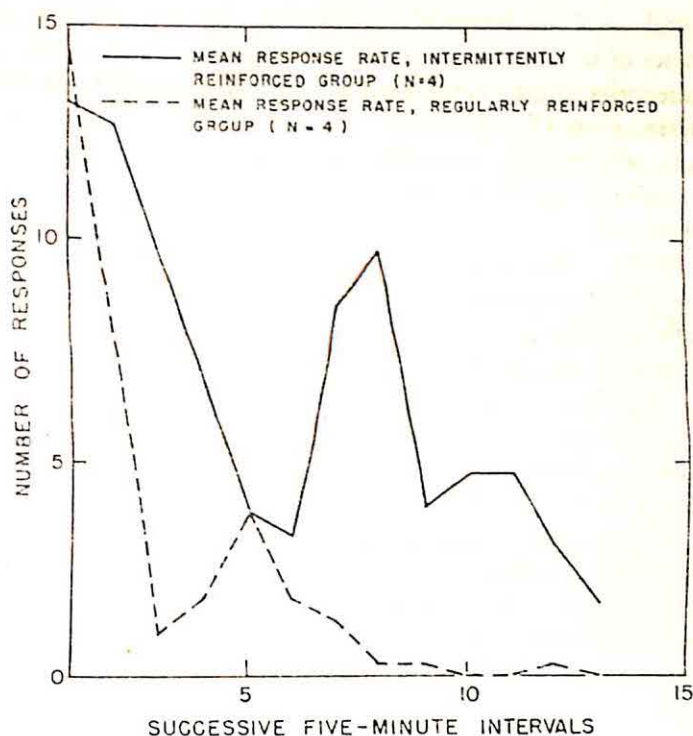


FIGURE 2—Noncumulative curves showing mean rates of smiling response during extinction period.

or right more than a few degrees. The "refusal" to fixate persisted even under these conditions; S's eyes then turned toward the ceiling. When withdrawal of reinforcement is conceptualized as frustration-producing (1), this persistent nonfixation may be regarded as an avoidance response that is elicited by continued frustration of the original (approach) response, is incompatible with the original response, and is reinforced by repeated escape from the frustrating situation.

The Relation of Protest to Smiling

As mentioned above, one of the criteria for subject selection was the frequency and intensity of crying. This criterion necessitated continued notations of such behavior during the operant period and first few conditioning intervals. Even after this point, however, E continued to record the incidence of crying, although at the time there seemed to be no particular reason for continuing to collect these data.

In recording crying, one of three types of notation was made, according to the intensity of response. In decreasing order of intensity, the notations were: (a) crying; (b) intense fussing: the same type of muscular and vocal involvement as in crying, but tears appeared only at the corners of eyes and did not course down the cheeks or temples; (c) fussing: considerable

muscular and vocal involvement, but less than for the first two categories; no appearance of tears. In the presentation of data and discussions to follow, all three categories will be considered as one and referred to by the generic response term, *protest*.⁷

The data concerning protest show that during the conditioning period, as rate of smiling increased, rate of protest decreased—or more properly, protest extinguished with the counter-conditioning of smiling. Specifically, there is a perfect rank correlation between number of trials taken to extinguish the protest response and number of trials to conditioning criterion on the smile response. Similarly, during the extinction period, as rate of smiling decreased, rate of protest increased. In this case, the rank order correlation (τ) between rates of emission of the two responses from the beginning of extinction up to the first interval at zero rate of response, is $-.69$, significant at the $.02$ level by a two-tailed test. (First appearance of zero response rate was chosen as the referent point for the extinction period because it was the closest counterpart of the referent point for the conditioning period, i.e., criterion or beginning of maximum response rate.)

To express these results more generally, for all Ss combined, the ratio of protests to smiles during the conditioning period was $1:6.5$ preceding criterion, and $1:276$ following criterion. During the extinction period, the corresponding ratios were $1:40.5$ (preceding first interval at zero rate of response) and $1:2.7$ (during and following first interval at zero rate of response). Rates of protest response during operant, conditioning, and extinction periods are shown in Figure 3.

One might conceivably object to any interpretation of these correlations on the grounds that they merely reflect the fact that two such responses are mutually exclusive behaviors: an infant cannot protest and smile at the same time. Although this overlooks the important point that all Ss *changed* in the frequency of emission of both responses, there is a direct answer to such an objection in that the precriterion data show, for every S, both part and whole intervals during which neither protest nor smiling occurred at all.

There is a good deal of similarity between these findings and results obtained in a study by Estes (3), in which rats were reinforced for one instrumental response while a previously reinforced, competing response was being extinguished. Estes states that, "These results seem to lend some support to the view that the 'learning' of any one response involves the concurrent extinction of others and that the amount of initial acceleration

⁷ It should be noted that interjudge reliability in discriminating protest from nonprotest was not determined. However, as indirect support for the objectivity or reproducibility of these data, the following points should be considered. First, this was a simple discrimination to make; even responses of lowest intensity represented a marked contrast to the infants' typical placid behavior. Second, the total numbers of protests for the two groups were approximately equal. Third, the orderliness of these data was not observed until the end of the experiment, when the results were actually tabulated.

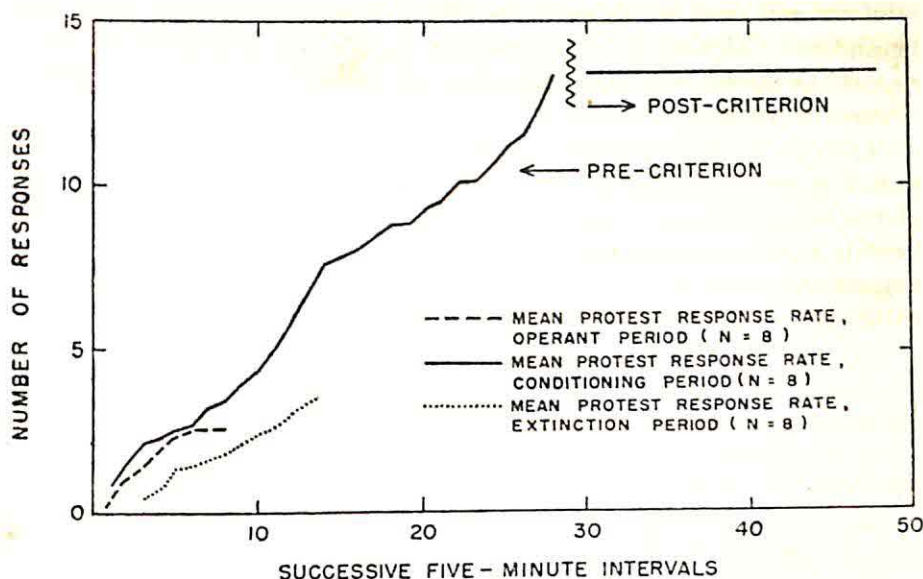


FIGURE 3—Cumulative curves showing mean rates of protest response during operant, conditioning, and extinction periods.

in a learning curve is determined to an important extent by the relative initial strengths of all behaviors which may occur in the experimental situation" (3, p. 204).

In the present case, the most important determiner of rate of acquisition of the smiling response appears to be the initial strength of the functionally equivalent, competing protest response. The main difference between procedures is that in Estes' study, the conflicting response was both reinforced and extinguished as an integral part of the experiment, while in the present study only the extinction of the competing response took place during the experiment. Its establishment had already taken place prior to the beginning of the experiment—most probably via reinforcement by maternal care.

Perhaps the most appropriate way of conceptualizing these results is in terms of a habit family hierarchy of two responses for which the initially differing habit strengths were first reversed by selectively reinforcing only the weaker response, and then reversed again by extinguishing that response, allowing for recovery of the first. This leads to the general proposition that rate of acquisition and extinction is a function of the initial discrepancy in habit strength between competing responses.

SUMMARY

This investigation was concerned with the instrumental conditioning of a social response (smiling) in infants. Ss were two groups of four infants each; one group was maintained on a conditioning schedule of intermittent reinforcement and the other on a schedule of regular reinforcement. The

CHILD DEVELOPMENT

divided into 39 pairs with one pupil of each pair having attended a nursery school while the other member of each pair had not.

In order to help equate these pairings the two pupils put together within each class were similar in sex, fathers' occupational level, and number of siblings in the family. Classification of fathers' occupations was based on the Revised Scale for Rating Occupations as given in Warner, Meeker, and Eells (14).

All of the children in the above listed grades responded to a one-criterion sociometric test involving several choices per pupil for preferred playmates. Total scores were calculated for each pupil on the basis of this measurement.

The teachers in four of the classes, including kindergarten through third grade, rated all the pupils included in their respective classes on the Winnetka Scale for Rating School Behavior and Attitudes (now out-of-print). This scale has twelve subdivisions, but the ratings obtained on all of these are summarized into five major trait-categories as follows: Co-operation, Social Consciousness, Emotional Adjustment, Leadership, and Responsibility. The teachers did not know the purpose of this study when they made their ratings.

Findings

The difference between the mean sociometric scores of the nursery and non-nursery school pupils was found by *t* test to be significant at the .02 level of confidence.

In the four classes in which teacher ratings were obtained there were 23 of the 39 pairings described above. When the mean score of the Winnetka Scale in each of the five major categories for the 23 pupils who had attended nursery school was compared with that for the 23 non-nursery school pupils, only one of the differences was found to be statistically reliable, this at the 5 per cent level for the trait designated Social Consciousness.

Summary

These findings clearly show that the pupils who had attended nursery school had a reliable advantage over the non-nursery school pupils in receiving positive choices from their classmates, but they showed a reliable advantage in only one out of five traits from the standpoint of teacher ratings on the Winnetka Scale.

SECOND STUDY

Methods and Subjects

The second study was a master's thesis completed at North Texas State College by Nicholson (10). Her subjects were 402 children in grades one through six in a public elementary school in a predominantly lower-middle class neighborhood in Dallas, Texas.

The first step in this study was to send a questionnaire to all the parents of all the pupils in this school asking them to supply information on whether or not their children had attended a nursery school or a kindergarten, or both. Of nearly 1000 questionnaires sent out, approximately 900 were returned. From these returns it was possible to select 14 classes (out of 32 classes available in the school) in which approximately 50 per cent of the pupils had had some type of preschool training. There were two classes on each of four grade levels and three classes on each of the other two grade levels.

Of the 402 pupils included in this study a little over 200 had had no formal preschool training. Of the nearly 200 who had had some type of preschool experience, approximately one-fourth had attended a nursery school, about one-half had attended a kindergarten, and the remaining one-fourth had attended both a nursery school and a kindergarten.

Classroom social adjustment was measured by sociometric testing and by teacher nominations.

In each of the 14 classes the pupils made four choices for others in their respective rooms with whom they would most like to eat lunch, and also four choices for the pupils they would select as candidates for the "Best Citizen of the Year" award—an annual custom in this school. These two criteria were used because they were both appropriate to all six grade levels in this school. Each child's total choice-status was based on his combined score from both of these criteria.

Teacher nominations were obtained by having each teacher list the upper one-fourth of pupils in her class who were regarded by her as making the most effective and satisfactory social adjustments within her classroom group; also, each teacher listed those whom she considered to be in the lowest fourth in reference to classroom social adjustments. All these listings were made before the teachers knew anything about the purposes of the investigation.

Findings

One of the major hypotheses which was tested in this study was whether the factor of preschool attendance or nonattendance was related to being in the upper or the lower halves of the sociometric distributions of the respective grade groupings. In order to have sufficient cases for statistical treatment the several classes on each grade level were combined, but all grade-level groupings were treated separately. The number of cases on the six grade levels ranged from 49 to 85 with a median of 67.

Using four-fold contingency tables each category was then compared with every other category. Because of the small number of cases in some cells an allowance for discontinuity was made by applying Yates' correction in computing the chi squares. When these data were assembled, it was found that none of the chi squares was statistically significant. These

findings obviously do not support the hypothesis that there was a significant relationship between choice-status, as measured in this research, and preschool attendance or nonattendance.

A second hypothesis tested in this investigation was that the teachers' listings of pupils in the upper and lower fourths from the standpoint of classroom social adjustments, as described above, would show a reliable relationship to the fact of preschool attendance or nonattendance. This hypothesis was also tested by chi square, using Yates' correction. None of these relationships proved to be even close to statistical significance.

A third hypothesis bearing on the carry-over value of preschool training was tested by forming matched groups. On each grade level those pupils who had attended some type of preschool were matched with nonattending pupils in regard to the following factors: sex, age, fathers' occupational classification, and IQs obtained from a group-administered intelligence test. On the basis of this matching the following number of pairs was obtained on each grade level from the first through the sixth: 23, 33, 34, 24, 20, 28. When the t test was applied to each of these experimental and control groupings, as they were related to sociometric choice-status on each grade level, none of the differences was found to be statistically significant. Only one of them even approached the .05 level of confidence.

Summary

It is clear that these findings provide no evidence that those pupils who had attended some type of preschool had, as a group, any advantage in personal-social behavior over those who had had no preschool training, either from the standpoint of acceptability by classroom peers or on the basis of teacher evaluations.

THIRD STUDY

Methods and Subjects

The third study was concerned with four sixth grade classes in the Denton, Texas, public schools. The pupils in each of these four sixth grades responded to a two-criteria sociometric test involving choices for preferred playmates and for desired associates in classroom group work. Also, each of the four teachers listed the names of those pupils whom she considered to be in the upper fourth of her class in "over-all good adjustment to the classroom social situation," and those she considered to be in the lowest fourth of her group on the same criterion.

Findings

In three of the four sixth grades approximately 35 per cent of the pupils had attended either nursery school or kindergarten. In these classes the average number of choices received on the two sociometric criteria combined was very nearly the same for the preschool and the non-preschool

groups. The actual difference was less than one point in each of the three classes.

In the remaining sixth grade 80 per cent of the pupils had attended either nursery school or kindergarten, or both. In this class the preschool pupils received an average sociometric score which was nearly twice as large as that received by the non-preschool pupils. This is obviously a marked advantage but probably should not be used as a basis for generalization because of the very unbalanced proportions of the two groups being compared in this class.

If the analysis of teacher placements into upper and lower fourths, as described above, is confined to the three classes in which the number attending some type of preschool amounted to approximately 35 per cent of each class, the conclusion is reached that these placements were unrelated to the factor of preschool attendance or nonattendance. Both the upper and lower fourths in each class included preschool and non-preschool pupils in numbers which were very close to what would be expected from their proportional representation in their respective groups.

Considering all four of the sixth grades, an effort was made to see if those pupils who had attended *both* nursery school and kindergarten would show up better on the sociometric testing and the teacher placements than those who had not attended either type of preschool. In all four classes combined there were 20 pupils who met this condition. An inspection of their average choice-status in their respective classes and their listings on the teacher placements failed to show any advantage for these pupils except in the room mentioned above in which only 20 per cent of the children had had no preschool training.

An effort was made to isolate the influence of the factor of preschool vs. non-preschool attendance by comparing these two groups in each class on the basis of IQs, fathers' occupational classification, and family size. The IQs were obtained by administering the California Short-Form Test of Mental Maturity. The fathers' occupational classifications were based on the same scale used in the first study. Family size was based on the number of siblings within six years of the age of each pupil being studied.

In none of these three variables was the difference between the preschool and the non-preschool groupings large enough to warrant statistical treatment. All the average IQ differences were less than seven points, and in three of the four classes the differences were only two or three points. All the differences between average scores for fathers' occupational classifications were less than one point with the exception of the class in which only 20 per cent of the pupils had not gone to preschool. In this class those who had attended preschool possessed an obvious advantage in fathers' occupational level.

In regard to family size the means and medians for all the groups were very similar, the largest single difference being less than one-half of a point.

CHILD DEVELOPMENT

Summary

From the foregoing data it is clear that the over-all picture is one of no significant differences between the preschool and non-preschool pupils in social adjustments in the sixth grades studied.

REVIEW OF RELATED RESEARCH

Other studies which have been designed to evaluate preschool education, particularly the nursery school, have shown some definite advantages for this type of education, although these advantages have not always been very outstanding or consistent on various types of measurements.

Kawin and Hoefer (7) have reported an investigation involving comparisons between 22 children who attended a nursery school and 22 who did not attend. No differences were found in physical characteristics or in degree of mental growth, but some differences were found in favor of the nursery school pupils in eliminating certain undesirable habits and in developing more independence of adults.

In a more extensive study Hattwick (5) compared ratings obtained by three nursery school teachers on 106 pupils who had been in nursery school attendance for nine months with an equal number who had been in attendance in the same schools only six weeks and who were matched with the first group on age, sex, nationality and race, and economic level. On most items involving social behavior and routine adjustments (but not emotional traits) the more experienced nursery school pupils were reliably superior to those who had been in attendance only six weeks.

Other similar studies by Joel (6), Parten (8), and Walsh (13) have also shown some beneficial results in various kinds of personal and social behavior from nursery school attendance.

On the other hand, Andrus and Horowitz (1) found little evidence that length of time spent in a nursery school increased feelings of security when these feelings were measured by three adult raters. The authors, however, were quite critical of the rating scale which they used. In a similar vein, Cushing (3) found "no striking differences" in personality trait ratings as given by nine kindergarten teachers to 33 pupils who had previously attended a nursery school as compared with a corresponding group who had not attended. Likewise, Phillips (9) found no advantage among first graders on the California Test of Personality for pupils who had had kindergarten training over those not having had such training.

DISCUSSION

One question which would naturally be raised in a report of this kind is in regard to the *quality-level* of the preschool training which any particular group of pupils receives. The only basis available for dealing with this point in reference to the three studies which constitute the main

body of this report is the proportion of subjects that attended a nursery school or a kindergarten sponsored by one of the two state colleges in Denton, the assumption being that these preschools are probably representative of the best kind of training available.

In the first study conducted by Angell nearly all the pupils had had their training in a college-sponsored preschool; in the second study probably none of the pupils had attended a college-sponsored preschool; while in the third study about 40 per cent of the sixth graders who had had preschool training had attended one of the college-sponsored nursery schools or kindergartens.

The one finding which showed the most clear-cut advantage for the preschool pupils was from Angell's study in which a very high percentage of the pupils had attended a college-sponsored preschool, and was based on sociometric data. However, teacher ratings in four of the six classes in this study showed only a small advantage for those who had attended these college-sponsored preschools.

Data from the four sixth grades in the third study do not lend support to the idea that attendance at a preschool of high quality may be reflected in better personal-social adjustment over a period of five or six years; but these data have only limited relevance to this point since not more than 40 per cent of the preschoolers in these classes had attended a college-sponsored nursery school or kindergarten.

One experimental study by Thompson (11) is especially pertinent to the matter of quality-level since his evidence showed that nursery school pupils whose teachers manifested toward them a large amount of warm friendly and helpful relationships made much more progress in various aspects of personal-social behavior than did a similar group of nursery school pupils whose teachers were considerate but who were somewhat detached and gave help only when it was specifically requested. These findings suggest that one of the essential characteristics of a preschool of high-quality level is a large amount of personal-social interaction not only between the pupils but also between the pupils and their teachers. Unfortunately, information is lacking on the extent to which this kind of relationship was present in the great majority of preschools that have been studied.

Another factor or condition pertinent to the present discussion is the extent to which the benefits of nursery school training can be shown to carry over into subsequent years, a determination of which was the main purpose of the three studies reported here. It would seem that this type of information is especially needed to evaluate preschool education since most of the studies previously reported have been confined to evaluating preschool experiences only during the time the preschool has been in session, with no subsequent follow-up. This is true of all of the studies reviewed or referred to above with the exception of the ones by Cushing and by Phillips and neither of these showed any advantage for preschool training. Without some efforts to trace the effects of preschool education

into subsequent years there is not an adequate basis for evaluating the contributions of special training as compared to simply general development.

One extensive investigation which has a bearing on this topic has been reported by Van Alstyne and Hattwick (12). They conducted a follow-up study into the elementary school grades of 165 children who had had nursery school training. One of the outstanding findings was a very strong tendency for those who were classified as well adjusted or as poorly adjusted in the nursery school to be likewise classified in their subsequent elementary school classes. The authors state: "The whole trend of the results . . . shows how closely behavior in the nursery school parallels that of the upper grades" (p. 67).

This finding, together with a corresponding finding in a similar study by Fisher (4), emphasizes the persistence of a certain level of personal-social adequacy irrespective of group socialization experiences.

These results, together with the evidence presented above in the three studies from the North Texas area, lead to the conclusion that if early socialization experiences are going to possess significant carry-over values into subsequent years they probably will need to be of a particularly high quality-level in reference to interpersonal rapport between the pupils and their adult supervisors and also in regard to adequate provision for meeting the varying needs of individuals.

It seems likely that many educators and child psychologists have had too much naive faith in the adjustment or curative values of group socialization, without sufficient attention being paid to whether or not the activities engaged in are actually helping those who need some assistance as opposed to simply perpetuating an established social hierarchy.

If this is a valid statement, then it follows that nursery school and kindergarten teachers should become more critical of their programs, particularly from the standpoint of planning more exercises, both in the school and with parents, which are designed to help those who show obvious signs of personal and social maladjustments.

REFERENCES

1. ANDRUS, R., & HOROWITZ, E. L. The effect of nursery school training: insecurity feelings. *Child Developm.*, 1938, 9, 169-174.
2. ANGELL, DOROTHY B. Differences in social behavior between elementary school children who have attended nursery school and those who have not attended nursery school. Unpublished master's thesis, North Texas State College, 1954.
3. CUSHING, H. M. A tentative report on the influence of nursery school training upon kindergarten adjustments as reported by kindergarten teachers. *Child Developm.*, 1934, 5, 304-314.

4. FISHER, M. L. Measured differences between problem and non-problem children in a public school system. *J. educ. Sociol.*, 1934, 7, 353-364.
5. HATTWICK, B. W. The influence of nursery school attendance upon the behavior and personality of the pre-school child. *J. exp. Educ.*, 1937, 5, 180-190.
6. JOEL, W. The influence of nursery school education upon behavior maturity. *J. exp. Educ.*, 1939, 8, 164-165.
7. KAWIN, E., & HOEFER, C. *A comparative study of a nursery school versus a non-nursery school group*. Chicago: Univer. of Chicago Press, 1931.
8. PARTEN, M. B. Social participation among pre-school children. *J. abnorm. soc. Psychol.*, 1932-1933, 27, 243-269.
9. PHILLIPS, G. J. Kindergarten training as a factor in the social adjustment of a selected group of young children. Unpublished master's thesis, Eastern New Mexico Univer., 1953.
10. NICHOLSON, ERTIE L. The relative social development of children with pre-school background as opposed to those who lack such experiences. Unpublished master's thesis, North Texas State College, 1957.
11. THOMPSON, G. G. The social emotional development of pre-school children under two separate types of educational program. *Psychol. Monogr.*, 1944, 56, No. 5.
12. VAN ALSTYNE, D., & HATTWICK, L. A. A follow-up study of the behavior of nursery school children. *Child Developm.*, 1939, 10, 43-70.
13. WALSH, M. E. The relation of nursery school training to the development of certain personality traits. *Child Developm.*, 1931, 2, 72-73.
14. WARNER, W. L., MEEKER, M., & EELLS, K. *Social class in America*. Chicago: Science Research Associates, 1949.

A SUGGESTED REVISION OF NORMS FOR THE ROGERS TEST OF PERSONALITY ADJUSTMENT¹

LEE G. BURCHINAL, BRUCE GARDNER, and GLENN R. HAWKES
Child Development Research Laboratories
Iowa State College

Although it is an old test, the Rogers Test of Personality Adjustment is still used in child guidance and school guidance programs. Yet, there have been few subsequent publications relating to this test since Rogers' original report (5). Recently this test was used in a study based on a sample of fifth grade children. The results of this study provided an opportunity to determine whether the norms given by Rogers in 1931 are still adequate for use today.

METHODOLOGY

The data for the present study are based upon completed Rogers tests from 256 fifth grade children selected from rural areas and small towns (less than 10,000 population by the 1950 Census) in Iowa, Ohio, Kansas, and Wisconsin. Eight sampling points were selected for each state sample by a probability method. At each sample point, eight children were randomly selected from the fifth grade. All the children came from "whole" families, ones in which both parents were living together, and all the children had at least one brother or sister. Field work was completed during the 1954-1955 school year. Detailed descriptions of the sample design and the children's family characteristics have been reported elsewhere (1, 2, 3).

¹ Published as Journal Paper No. J-2963 of the Iowa Agricultural Experiment Station, Ames, Iowa, Project No. 1171, Home Economics Research. Acknowledgment is due Leone Kell of the Kansas Agricultural Experiment Station, Ruth Hoefflin of the Ohio Agricultural Experiment Station, and Helen Dawe of the Wisconsin Agricultural Experiment Station for their collaboration in this regional research project.

CHILD DEVELOPMENT

Since descriptions of the Rogers Test of Personality Adjustment are available (4; 5, pp. 354-356), only a brief discussion is given here. Four subscores, personal inferiority, social maladjustment, family relations, and day dreaming, are derived from the test. When 43 children were retested by Rogers after an interval of one month, subscore reliabilities ranged from .65 to .72. The total score reliability was .72. In the present study, 51 children were retested after a one-week interval, and subscore reliabilities varied from .67 to .77. The total score reliability was .71. The results of the several efforts made by Rogers to test the validity of the adjustment tests were accepted for the purposes of the present study.

Rogers developed a set of norms for use with his test based on a total of 167 children. In developing his norms, Rogers used clinical data as well as means and standard deviations. Taking these factors into account, Rogers divided the scores into three groups, high, average, and low levels of maladjustment, because he felt finer gradations were not warranted on the basis of his data. The extreme categories approximated quartile groupings of scores. For Rogers' norms, see Table XX (5, p. 86).

TABLE I
PERCENTAGE OF LOW, AVERAGE, AND HIGH SCORERS ACCORDING
TO ROGERS' NORMS

<i>Adjustment Scores</i>	SCORE CLASSIFICATION		
	<i>Low</i>	<i>Average</i>	<i>High</i>
Personal Inferiority	52	28	20
Social Maladjustment	12	36	52
Family Relationships	27	41	32
Day Dreaming	38	29	33
Total Score	23	51	25

THE FINDINGS

When the scores of 256 children tested in this study were classified according to Rogers' norms for the four subscores and the total scores, the results presented in Table I were obtained.

It should be noted that on the personality inferiority scale 52 per cent of the children were low scorers while on the social maladjustment exactly the same percentage were high scorers. It would seem that any set of norms that places over half of the cases into one or the other extreme scoring category is unrealistic. The percentage distributions for the family relationships scores approached what might have been expected, while the distribution of the total scores followed closely a quartile distribution. The distribution for the day dreaming scores was bimodal with the smallest frequency in the "average" category.

In order to test whether the frequencies for the percentages listed in Table 1 differ significantly from reasonable estimates of the distributions of the scores, an expected distribution for the scores was defined by using quartile groupings of scores. Except for the absence of clinical data, this method of defining low, average, and high scoring categories approximated a replication of the procedure followed by Rogers. Each of the extreme categories should have included approximately 64 cases while the average category should have included 128 cases. Chi square comparisons were used to test the departure of the observed from the expected frequencies. For simplicity, all the computation tables have been omitted.

The distribution for the total scores was the only one which did not depart significantly from the expected distribution ($.50 < p < .70$). The chi squares for the four subtests were significant ($p < .01$).

Since the observed frequencies for the classification of scores departed significantly from what might have been expected on all four of the sub-scores, a new set of norms based on the scores of the 256 children tested in this study was developed. Differences between boys' and girls' mean scores were tested first to determine if separate norms ought to be used for some of the scores. Mean scores and standard deviations for boys and girls and for the total sample for the five scores are given in Table 2.

TABLE 2
MEAN SCORES AND STANDARD DEVIATIONS FOR BOYS, GIRLS, AND
TOTAL SAMPLE ON ROGERS TEST

	Boys N=129		Girls N=127		TOTAL SAMPLE N=256	
	M	SD	M	SD	M	SD
Personal Inferiority . . .	10.81	3.95	12.43	4.99	11.61	4.57
Social Maladjustment . .	14.74	5.74	14.85	4.07	14.80	4.43
Family Relations	9.18	3.88	8.72	3.78	8.95	3.84
Day Dreaming	2.88	2.28	3.31	2.77	3.09	2.55
Total Score	37.60	8.00	39.17	8.39	38.38	8.23

Since the boys' and girls' means were significantly different on the personal inferiority scale, it seemed that separate norms should be considered for these scores; otherwise, similar norms were developed for boys and girls for the other four scores.

Suggested norms for the Rogers scores are listed in Table 3. The cutting points for the low, average, and high scoring groups were determined by taking the scores which came the closest to the dividing lines for the top and bottom quartile.

CHILD DEVELOPMENT

The modified norms differ considerably from Rogers' norms for the personal inferiority and social maladjustment dimensions of the adjustment test. Also, separate norms for boys and girls are suggested on the personality inferiority scores. Only small changes occurred in relation to the family relationships scores. Several changes were observed for the day dreaming scores. Rogers suggested separate norms for boys and girls on this dimension of his test. The present data suggested separate norms were not necessary. The new norms did not conform to either of the sets given by Rogers, but approximated the averages of the separate norms given by him. Finally, no changes were necessary for the total score norms.

TABLE 3
SUGGESTED NORMS FOR THE ROGERS TEST

<i>Adjustment Scores</i>	SCORE CLASSIFICATION		
	<i>Low</i>	<i>Average</i>	<i>High</i>
Personal Inferiority			
Boys	- 8	8-13	14+
Girls	- 9	9-15	16+
Social Maladjustment	- 12	12-17	18+
Family Relations	- 7	7-11	12+
Day Dreaming	- 2	2- 4	5+
Total Score	- 33	33-43	44+

SUMMARY

Data derived from the administration of the Rogers Test of Personality Adjustment to a sample of 256 fifth grade children provided a basis for evaluating the norms given by Rogers in 1931. Rogers' norms appeared to be inadequate for classifying children into low, average, and high scoring categories for all the subscores of the test. The norms for the total scores are apparently still adequate.

When it was apparent that Rogers' norms were not realistic for present use of the test, boys' and girls' mean scores were compared to determine if separate or similar norms should be used for them on the various Rogers scores. The means of the two groups were not significantly different except for the personal inferiority scores. Therefore, boys and girls were separately classified for the development of norms for the personal inferiority scores. The total sample of all children was used for the other scores. Scores corresponding to the cutting points for the top and bottom quartiles were used to define the low, average, and high quartiles. On this basis, a new set of norms was proposed.

REFERENCES

1. BURCHINAL, L. G. The relations of parental acceptance to adjustment of children. Unpublished doctor's dissertation, Ohio State Univer., 1956.
2. BURCHINAL, L. G., HAWKES, G. R., & GARDNER, B. The relationship of parental acceptance to the adjustment of children. *Child Develpm.*, 1957, 28, 65-77.
3. HAWKES, GLENN R., BURCHINAL, L. G., GARDNER, B., & PORTER, B. M. Parents' acceptance of their children. *J. Home Econ.*, 1956, 48, 195-200.
4. REMMERS, H. H., & GAGE, N. L. *Educational measurement and evaluation*. New York: Harper, 1943.
5. ROGERS, C. *Measuring personality adjustment in children nine to thirteen years of age*. New York: Teachers College, Columbia Univer., 1931.

RESEARCH ON THE EFFECTS OF NON-CONTINUOUS MOTHERING¹

DAMARIS PEASE *and* D. BRUCE GARDNER
Child Development Research Laboratories
Iowa State College

That there is some significant relationship between the development of the infant and the nature of his early mothering experiences would not appear to be open to serious question at this point. The negative effects of inadequate mothering in the case of the severely deprived infant have been repeatedly observed by a number of research workers (2, 5, 9, 10, 12, 15, 16, 17, 19, 20).

Bowlby (5), in a summary of such research, considers it highly significant that different research workers, operating independently and to some degree in ignorance of each others' efforts, arrived at essentially similar conclusions on the adverse effects of mother-child separation.

A clinical description of the deprived infant, which for all its brevity is remarkably comprehensive, is presented by Bakwin:

Infants under 6 months of age who have been in an institution for some time present a well-defined picture. The outstanding features are listlessness, emaciation and pallor, relative immobility, quietness, unresponsiveness to stimuli like a smile or a coo, indifferent appetite, failure to gain weight properly despite the ingestion of diets which, in the home, are entirely adequate, frequent stools, poor sleep, an appearance of unhappiness, proneness to febrile episodes, absence of sucking habits (3, p. 512).

Other writers (16, 20) have stressed the finding that institutional children in a physically sterile setting are more susceptible to colds and epidemics of infectious diseases than are children who have the psychological support of a family atmosphere.

Bowlby lists the typical characteristics of children deprived in infancy, when observed later in childhood, as follows: "... superficial relationships;

¹ Journal Paper No. J-3281 of the Iowa Agricultural and Home Economics Experiment Station, Ames, Iowa, Project No. 1299. Acknowledgment is due the Elizabeth McCormick Memorial Fund for partial financial support.

no real feeling—no capacity to care for people or to make true friends; an inaccessibility, exasperating to those trying to help; no emotional response to situations where it is normal—a curious lack of concern; deceit and evasion, often pointless; stealing; lack of concentration at school” (5, p. 31).

One is impressed, in reading such descriptions, that this wide range of symptoms represents not merely an isolated effect on some specific segment of personality, but a vast complex of responses of the total organism to an inadequate environment. It is conceivable that this complex of organismic responses, as the deprived individual moves forward over the time dimension, has its own pattern of lawful and orderly development, which we are only beginning to understand. Bowlby and Robertson, at least, are willing to assert that, “The general proposition that separation in the early years of life can have permanent adverse effects now belongs to history” (7, p. 123). They go on to say that we still need to know how these effects come about.

Goldfarb (15) presents a rather striking picture of the differences in behavior between children who were institutionalized as infants and children who were reared in foster homes. His comparisons were made between two small groups of children between the ages of 10 and 14, with respect to intelligence, ability to conceptualize, school achievement, social maturity, speech, response to frustration, and capacity for relationships. He concluded that the institutional group was significantly inferior to the control group in all of these characteristics. Comparison of Rorschach responses for the two groups led Goldfarb (12) to characterize the perceptions of the institutional children as looser, less well organized, confabulatory, and more arbitrary. He found lack of control over emotional responses for the institutional children, with evidence of diminished drive to social conformity.

To the extent that we can accept and generalize the findings of such research we might conclude that extreme and severe deprivation in early childhood initiates a progressive and irreversible developmental pattern. Further it would seem that the total organism is involved in a complex set of relationships in the progressive developmental pattern, and that the deprived child is especially characterized by reduced ability to utilize his personal and environmental resources, and lessened capacity for healthy affectional relationships with others.

The research on which such conclusions might be based has dealt largely with institutionalized infants and children, and has tended to make deprivation operationally synonymous with institutionalization. It becomes easy to think of the mother-child relationship as if it were a dichotomous variable, with deprivation either occurring or not occurring in any given situation. There is obviously no basis for making such an assumption; indeed it is probably safe to assume just the opposite, that there is infinite and continuous variation from some condition representing a most desirable mother-child relationship to a condition of complete deprivation. Although even these extremes are not precisely defined in the literature, it is likely

that considerable agreement could be found among professional people as to what constitutes the essential features of these conditions. But between these two relatively obvious extremes, it is less clear as to what constitutes intermediate amounts, kinds, and qualities of mothering, and what the effects of these relative degrees or conditions of mothering might be. Furthermore, it is a questionable assumption that only one continuum is involved. The complexity of factors suggests that the mother-child relationship has many dimensions, including that of time. That is, the nature of the conditions constituting adequate mother-child relationships may vary considerably with the age of the child. Thus, mothering itself, if defined operationally, must take into account the varying demands of a relationship which changes with time. Similarly, an operational definition of deprivation in all its infinite variations would necessarily take into consideration the notion that a given set of conditions might constitute varying degrees of deprivation at different times.

Goldfarb (15), in attempting to outline the differences between adequate and inadequate mother-child relationships, points out that in the ordinary family (a) the child has *specific* adults (parents) who are warm and loving; (b) the mother influence is *continuous*, in terms of daily routine and total life-span; (c) the mother is the source of safety and need-gratification to the child; (d) the child is given *constant stimulation* from his world of persons and objects; and (e) there is a reciprocal relationship between parent and child. The absence of these elements is considered maternal deprivation.

Even in the event of a clear-cut separation of the infant from its mother, there are obviously many conditions which may modify the effects of the experience on the child. Ainsworth and Bowlby comment on this:

Both theoretical considerations and research to date suggest that separation experiences occurring under different conditions lead to effects which may differ qualitatively and/or quantitatively from one another. Certain of the conditions which seem likely to affect the outcome of separation . . . (include) . . . the state of psycho-social development and the strength of the dependency relationship which have been achieved by the time of separation and the nature of the relationship with substitute figures during the separation. Other conditions are: the degree of stability and security of the pre-separation dependency relationship with the mother; the length of the separation; the degree of continuity of the mother-child relationship during separation (for example through visiting); the nature of other relationships before and after separation, particularly the father-child relationship; the degree of previous familiarity with the separation environment and substitute figures, if any; the degree of frustration of other needs such as perceptual and activity needs during separation; and the reasons for separation and the child's understanding of them (1, p. 109).

The authors go on to suggest that we need to learn more about the role of such variables in modifying the effects of separation experience on the child. It is evident that there are a good many things we do not yet know, not only about the effects of separation and extremes of deprivation,

CHILD DEVELOPMENT

but about the basic elements of mother-child relationships which determine success, if the criterion for successful mothering is the optimum development of the child. The research of Dennis (8), which led to conclusions not entirely consistent with other research cited here, raises questions about the ostensibly deprived child whose development does not seem to be adversely affected by his experience. Stone points up our lack of understanding on such matters and urges cooperative research efforts:

What, precisely, are the significant ingredients in the mothering relationship? Is it a specifically humanizing process, or to what extent is it essential to mammalian development in general? . . . What mothering practices are particularly harmful at which stages of development, and at which later stages of development will the damage or the benefit be evident? Here is opportunity for physician, experimental psychologist, clinical psychologist, case worker, and sociologist to join hands in order to arrive at some of the very fundamentals of the influences which shape and misshape human personality (23, p. 18).

Of special interest to the writers in their current research pertaining to non-continuous mothering has been the emphasis placed by Bowlby on the continuity of the mother-child relationship. At one point, in commenting on mental health, Bowlby defines mental ill-health as ". . . a derangement of the capacity to make co-operative relationships with other human beings (6, p. 117). Elsewhere Bowlby states: ". . . What is believed to be essential for mental health is that the infant and young child should experience a warm, intimate, and continuous relationship with his mother (or permanent mother-substitute) in which both find satisfaction and enjoyment" (5, p. 11).

Thus, it would appear that, to Bowlby at least, any mother-child situation in which this warm, intimate, *and continuous* relationship does not exist can be considered a state of maternal deprivation.

The research to be described here, currently in progress, represents one attempt to explore the continuity variable, and to determine what, if any, relationship might exist between the growth of the child and the continuity of his mothering experience in early infancy.

At certain colleges and universities the practice, in departments of home economics, has been to bring an infant into a home management house for a period of time to provide the students with an educational experience in the care of infants. Critics of this practice have argued that it constitutes a kind of maternal deprivation for the infant, since a number of women students, usually in rotation, have primary responsibility for the actual care of the infant. Thus, he is deprived of an experience with a single, continuous mother figure. A review of the literature fails to reveal any definitive statements of the effects on the child of being subjected to such a discontinuity of mothering.

Child Development and Home Economics personnel at Iowa State College have been interested in this question for a number of years. In

1955, in cooperation with the Iowa Children's Home Society, they sought and obtained financial support from the Elizabeth McCormick Memorial Fund to conduct research on the effects of home management house experience on the personality development of infants.

At the six home management house units currently operating at Iowa State College, the practice is to bring the infant into the house for a period of one college quarter, beginning usually when the infant is less than one month in age. Each house is equipped to accommodate a group of eight students and the house adviser. The students live in the house for half of one quarter, or about five and one-half weeks. Each student is given primary responsibility for the care of the infant over a four- or five-day period. With a new group of students entering the house at mid-quarter, the infant is subjected to a noncontinuous mothering experience with as many as 16 different mother figures over a period of approximately three months. In addition, there is the discontinuity associated with moving the infant either from the hospital or from a foster home to the home management house, and from the home management house to either a foster home or an adoptive home.

In research currently in progress, the experimental group consists of approximately 40 infants who have lived in or will live in a home management house for a period of one college quarter, beginning in January, 1956. Two other groups of infants, representing two different classes of experimental control, are also being observed. The first of these is a group of infants, matched for age and sex with the experimental group, who are subjected to foster home care for a period of time during early infancy, followed normally by placement in a private adoptive home. The second of these control groups consists of infants born into complete families who remain with their natural parents and experience no unusual mothering situation. Infants in this group are also matched for age and sex with the infants in the experimental group. For convenience, the three groups have been identified as follows: Group A—home management house infants, Group B—infants subjected to foster home care, Group C—infants subjected only to "normal" mothering in the homes of their natural parents.

Infants comprising Groups A and B are brought into the research through the services of the Iowa Children's Home Society. They are children of unmarried mothers. Efforts are made to control any selective factors which might operate to determine which infants are placed in Group A and which in Group B. Thus, it may reasonably be assumed that no important differences exist between the two groups at the time of entry into the research project. Infants in Group C are brought into the research through the cooperation of a pediatric clinic in Des Moines, Iowa.

All infants in the three groups are given an initial testing shortly after birth, one purpose of which is to determine whether they can be accepted as research subjects. Infants with obvious congenital defects such as spastic

paralysis, congenital malformations of limbs or body organs, micro- or hydrocephaly, or serious sensory defects, are excluded. Premature infants are also excluded.

At three months of age each infant is given a complete test battery, which is then administered again at each of the following ages: 6 months, 9 months, 12 months, 18 months, and 24 months. This test battery is given as nearly as possible to the exact day representing each of these ages. However, for practical reasons and difficulties of scheduling, a period of time from two weeks prior to that day to two weeks following that day has been allowed.

The kinds of data being collected on the infants may be classified as follows: (a) medical, (b) psychometric, and (c) anthropometric. The medical data consist of the birth history and reports made regularly by the examining physician. The psychometric data are derived through administration of the Gesell Developmental Schedule, the Cattell Infant Intelligence Scale, and the California Infant Scale of Motor Development. In addition, observational data representing the impressions of the psychological examiners are recorded during the examination period. The development of basic reflex patterns is also noted and recorded. At the time of the 12-months examination a brief motion-picture sequence is made of the infant in a relatively standardized situation. This sequence is taken immediately following the completion of the anthropometric measurements, while the infant is being dressed. This provides a visual record of each subject at 12 months of age, made under comparable conditions for all infants.

Anthropometric measurements recorded at each of the examination ages are: head circumference, chest circumference, calf circumference, bi-iliac width, thickness of subcutaneous fatty tissue in calf of left leg, over-all length, and weight.

The medical data are obtained by a physician within a few days of the psychological examination. Psychological data are obtained by a team of examiners, usually two persons working together. Anthropometric measurements are taken by a nurse and an assistant, both of whom have received instruction in the techniques of taking physical measurements of infants.

The psychological and anthropometric measurements have been made, in the majority of cases, either in a testing room at the Iowa Children's Home Society in Des Moines, or in a testing room of the Child Development Laboratories at Iowa State College. For the Group C infants the initial examination has been conducted in one of the examining rooms of a pediatric clinic in Des Moines; for the subsequent tests the Group C infants are brought by their mothers to the Iowa Children's Home Society. A few infants in Groups A and B, who are generally adopted before the 6-months test, have sometimes been examined in their private adoptive homes if for some reason the parents could not bring them to one of the testing centers. In such cases the examining teams have attempted to reproduce as many of the testing room conditions as the home situation would

permit. The administration of the complete test battery, excluding the medical examination which is given separately by the physician, requires about one hour, depending on the age of the infant, the ease with which rapport is developed, and other miscellaneous factors.

At the time of this writing, there are 101 subjects in the noncontinuous mothering research project, about evenly distributed among the three groups. A preliminary comparison of 36 infants in Group C with 60 infants in the combined Groups A and B indicates that the home management house and foster home babies are in no way inferior to the natural home babies at the time of induction into the research.

As the data are analyzed, reports will be made which describe more fully the comparative status of Groups A, B, and C at the time of initial examination, as well as the relative progress of each of the three groups during the first two years of life. It is to be hoped that follow-up studies of these same subjects will permit investigation of the long-range effects, if any, of the kind of discontinuous mothering described here.

REFERENCES

1. AINSWORTH, M. D., & BOWLBY, J. Research strategy in the study of mother-child separation. *Courrier*, 1954, 4, 105-130.
2. AUBRY, J. The case of Monique. In K. Soddy (Ed.), *Mental health and infant development*. Vol. I. New York: Basic Books, 1956. Pp. 125-130.
3. BAKWIN, H. Emotional deprivation in infants. *J. Pediat.*, 1949, 35, 512-521.
4. BENDER, L. *Aggression, hostility and anxiety in children*. Springfield, Ill.: Charles C. Thomas, 1953.
5. BOWLBY, J. *Maternal care and mental health*. Geneva: Monogr. World Health Organization, 1952.
6. BOWLBY, J. Mother-child separation. In K. Soddy (Ed.), *Mental health and infant development*. Vol. I. New York: Basic Books, 1956. Pp. 117-122.
7. BOWLBY, J., & ROBERTSON, J. A two-year-old goes to hospital. In K. Soddy (Ed.), *Mental health and infant development*. Vol. I. New York: Basic Books, 1956. Pp. 123-124.
8. DENNIS, W. Infant development under conditions of restricted practice and of minimum social stimulation: A preliminary report. *J. genet. Psychol.*, 1938, 53, 149-157.
9. FREUD, ANNA, & BURLINGHAM, DOROTHY. *Infants without families*. New York: International Universities Press, 1944.
10. GOLDFARB, W. Infant rearing and problem behavior. *Amer. J. Orthopsychiat.*, 1943, 13, 249-265.

11. GOLDFARB, W. Infant rearing as a factor in foster home replacement. *Amer. J. Orthopsychiat.*, 1944, 14, 162-166.
12. GOLDFARB, W. Effects of early institutional care on adolescent personality: Rorschach data. *Amer. J. Orthopsychiat.*, 1944, 14, 441-447.
13. GOLDFARB, W. Effects of psychological deprivation in infancy and subsequent stimulation. *Amer. J. Psychiat.*, 1945, 102, 18-33.
14. GOLDFARB, W. Variations in adolescent adjustment of institutionally reared children. *Amer. J. Orthopsychiat.*, 1947, 17, 449-457.
15. GOLDFARB, W. Emotional and intellectual consequences of psychologic deprivation in infancy: A revaluation. In P. H. Hoch and J. Zubin (Eds.), *Psychopathology of childhood*. New York: Grune & Stratton, 1955. Pp. 105-119.
16. HEINICKE, C. M. Some effects of separating two-year-old children from their parents. *Hum. Relat.*, 1956, 9, 105-176.
17. LOWREY, L. G. Personality distortion and early institutional care. *Amer. J. Orthopsychiat.*, 1940, 10, 576-585.
18. RHEINGOLD, HARRIET L. Mental and social development of infants in relation to the number of other infants in the boarding home. *Amer. J. Orthopsychiat.*, 1943, 13, 41-44.
19. RIPIN, ROWENA. A comparative study of the development of infants in an institution with those in homes of low socio-economic status. *Psychol. Bull.*, 1933, 30, 680-681.
20. SPITZ, R. A. Hospitalism. *Psychoanal. Study Child*, 1945, 1, 53-73.
21. SPITZ, R. A. Anaclitic depression. *Psychoanal. Study Child*, 1946, 2, 313-342.
22. SPITZ, R. A. Hospitalism. *Psychoanal. Study Child*, 1946, 2, 113-117.
23. STONE, L. J. A critique of studies of infant isolation. *Child Developm.*, 1954, 25, 9-19.

A NOTE ON THE SIMILAR EFFECTS OF LOW SOCIAL AVAILABILITY OF AN ADULT AND BRIEF SOCIAL DEPRIVATION ON YOUNG CHILDREN'S BEHAVIOR

JACOB L. GEWIRTZ
Laboratory of Psychology
National Institute of Mental Health

DONALD M. BAER
University of Washington

CHAYA H. ROTH
The University of Chicago

In two recent researches experimental conditions were devised to affect the behavior of young children for social reinforcers from adults. These conditions (which were implemented by different operational procedures in each case) appear to have a number of possibly critical formal characteristics in common. A comparison of them may prove fruitful.

PROCEDURES

In one study (2), children (Ss) ranging in age from 4-0 to 5-7 engaged individually in easel painting in the presence of an adult. They could emit attention-seeking responses to the adult under conditions of the apparent low and high availability of the adult as a source of social reinforcers. With one-half (28) of the Ss, the adult appeared continually available (High Availability); that is, he sat passively immediately behind the child, attending completely to him. The child was told at the outset that the adult would watch him paint. With the remaining 28 Ss, the adult sat at a desk farther away, apparently engrossed in paper work (Low Availability). Here the child was informed at the outset that the adult had work to complete at his desk, but that upon request he would be happy to supply all the materials the child needed. Under both conditions the adult responded briefly to any direct initiation of the child in a friendly and pleasant manner.

In the second study (4), 32 children in a similar age group (3-10 to 5-3) played a two-response marble game individually in the presence of an adult experimenter (E). By reinforcing the initially less frequent response with *approval* (i.e., by saying "Good," "Um-hmm," and the like), the relative frequency of the response could be increased. Each child played the game in the adult's presence on two occasions. One, defined as *Social Deprivation*, was preceded by a 20-minute period of social isolation. During

this deprivation period the child waited alone, in a bare room while the game was "repaired" by the adult in an adjoining room. On the other occasion, defined as *Non-deprivation*, the child played the game immediately upon his arrival from his play group.

Factorial control was exercised over the sex of the children and adults in both studies (and different Ss and Es were employed in each). That is, half of the subjects of each sex interacted individually with a man, the other half with a woman.

RESULTS

Low Availability of the adult and Social Deprivation both effected similar increases in the incidence of behaviors by the children for social reinforcers from the adults. The mean frequency of attention-seeking responses was greater under Low Availability than under High Availability. And, similarly, the mean relative frequency of behaviors for approval was reliably increased by the Deprivation condition, relative to the Non-deprivation condition.

In both studies a similar cross-sex pattern of behavior for a social reinforcer was observed: Children appeared to emit more reinforcer-oriented responses to adults of the opposite sex than to adults of the same sex as themselves. In both studies, this tendency was more marked for boys than for girls. Furthermore, this pattern uniformly qualified the experimental conditions determining the relative supply of these reinforcers to the children, in both studies. That is, there was a suggestion that whether or not Ss displayed more attention-seeking to adults of the opposite sex depended on whether the adult was under Low or High Availability. Similarly, the effect of Social Deprivation on what may be called approval-seeking was qualified by a reliable second-order interaction effect between the condition of Deprivation, the Sex of the Child, and the Sex of the Adult.

Figure 1 demonstrates the relative *change* in behavior for social reinforcers, which can be attributed to Low Availability in the one case and to Social Deprivation in the other, as reflected in the four conditions provided by the combinations of the variables Sex of Child and Sex of Adult. That is, increases in behavior for attention from the High Availability condition to the Low Availability condition are compared with increases in behavior for approval from Non-deprivation to Deprivation. The nature of the effect is seen to be remarkably similar in the two cases. If the four Availability mean differences are ranked from high to low, and compared to such rankings of the four Deprivation mean differences, we find that an identical pattern emerges: the rank order is Boys with Women, Girls with Men, Girls with Women, and Boys with Men. The Spearman rank correlation coefficient between these two patterns is $+1.00$ ($p < .05$, single-tail test). From this we may conclude that the concordance between the two patterns is reliable.

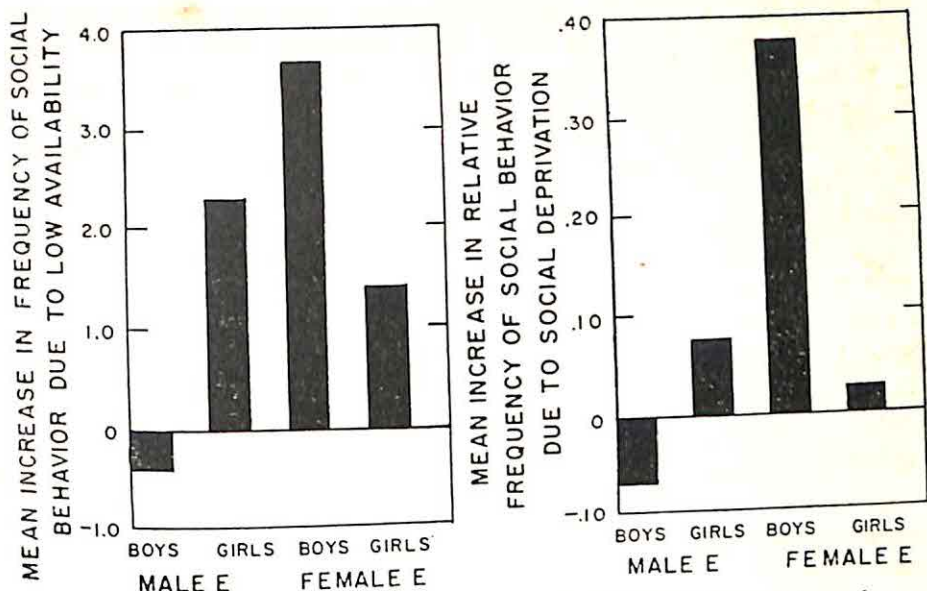


FIGURE 1—A comparison of the mean increases in behaviors for social reinforcers due to Low Adult Availability with those due to Social Deprivation, as reflected against the four Sex of *S* by Sex of *E* conditions.

COMPARISON

This uniform cross-sex qualification of the Availability and Deprivation variables might be referred to a number of rationales [summarized in (2)], the most prominent probably being Freud's Oedipus theory. However, the underlying nature of this interaction is less important here than the fact that it displays such a remarkably similar form under both experimental conditions. This similarity cannot readily be ascribed to chance; consequently, it presents a clear invitation to analyze the formal characteristics of the deprivation and availability conditions. If these conditions can increase behavior for what are presumed to be different social reinforcers in a similar manner under comparable conditions, it is reasonable to suspect that they share important characteristics.

An examination of these two experimental conditions bears out this expectation. Social Deprivation clearly involves a complete limitation of access to all sources of social reinforcement. A fair restatement of this would be that all reinforcing persons are rendered unavailable for a period of time. But the Availability conditions certainly appear to represent two points along precisely this dimension: Low Availability of the adult implies a more limited supply (although not a zero supply) of reinforcers from him than does High Availability.¹ In other words, both experimental pro-

¹ It should be noted that in the Deprivation condition, reinforcers were delivered *after* a period of zero supply of reinforcers; in the Low Availability condition, reinforcers were delivered *during* a period of apparently limited supply of reinforcers.

CHILD DEVELOPMENT

cedures appear to involve a simple dimension of the relative supply of social reinforcers.

Perhaps this dimension is logically similar to the deprivation and satiation operations controlling the primary appetitive drives (6). Recent research has shown that operations of this class may control reinforcers of no apparent biological significance, as, for example, the opportunity to make a brief observation response (1). The findings described here and in (5)² seem to suggest that apparently similar operations of deprivation and satiation may prove to be extendable to a wide class of social stimuli, and may constitute a simple convention defining social as well as appetitive drives (3). The research required for this problem presumably would focus initially upon the variety of operations which might be used to restrict or increase the relative supply of social reinforcers, and upon the possibility that these operations would show similar effects on a variety of social reinforcers.

REFERENCES

1. BUTLER, R. A. The effect of deprivation of visual incentives on visual exploration in monkeys. *J. comp. physiol. Psychol.*, 1957, 50, 177-179.
2. GEWIRTZ, J. L. Three determinants of attention-seeking in young children. *Monog. Soc. Res. Child Developm.*, 1954, 19, No. 2 (Serial No. 59).
3. GEWIRTZ, J. L. A program of research on the dimensions and antecedents of emotional dependence. *Child Developm.*, 1956, 27, 205-221.
4. GEWIRTZ, J. L., & BAER, D. M. The effect of brief social deprivation on behaviors for a social reinforcer. *J. abnorm. soc. Psychol.*, 1958, 56, 49-56.
5. GEWIRTZ, J. L., & BAER, D. M. Deprivation and satiation of social reinforcers as drive conditions. *J. abnorm. soc. Psychol.*, in press.
6. SKINNER, B. F. *Science and human behavior*. New York: Macmillan, 1953.

² A condition of *satiation* for social reinforcers (5) has produced effects consistent with this logic.

A SHORT FORM OF THE CHILDREN'S MANIFEST ANXIETY SCALE

NISSIM LEVY

Brown University

Recently Castaneda *et al.* have published a children's form (CMAS) of the Taylor manifest anxiety scale (1). In the process of carrying out a large scale standardization and item analysis study on subjects from two towns in Rhode Island, we found 10 items which showed no significant differences for either grade or sex.

It was the purpose of the present study to determine the relationship between this 10-item form of the CMAS and the original 42-item form. All protocols were rescored using the 10-item form. This was accomplished very simply since all of the protocols had previously been coded on IBM cards for the item-analytic study.

Since the Rhode Island data indicated that not only were there significant sex and grade differences but also a significant interaction for full-scale scores, separate Pearsonian correlation coefficients were computed for each of the six subgroups which comprised our total population of subjects. Table 1 presents these values.

It can be seen that the short-form scores do quite a respectable job of predicting the full-form scores. The correlations range in value from .84

TABLE I
SHORT FORM-FULL FORM CORRELATIONS FOR DIFFERENT GRADE LEVELS AND FOR BOYS AND GIRLS SEPARATELY

	G R A D E					
	4 T H		5 T H		6 T H	
	N	r	N	r	N	r
Boys	145	.84	129	.85	151	.88
Girls	130	.84	113	.84	185	.95

CHILD DEVELOPMENT

to .95. Approximately 70 to 90 per cent of the variance in the full-scale scores can be accounted for by the variance of the scores on the short form.

Another of the possible advantages of the short form lies in the fact that it is composed of items which do not seem to show significant sex or grade differences. One could, it would seem, more reasonably ignore Ss' sex or grade in the fourth to sixth grade range when using the short-form scale than when using the full scale.

For those interested in using the short form, the 10 items which comprise it are numbers 2, 9, 13, 15, 18, 26, 31, 43, 44, and 46 as presented in Castaneda *et al.* (1).

REFERENCES

1. CASTANEDA, A., MCCANDLESS, B. R., & PALERMO, D. S. The children's form of the manifest anxiety scale. *Child Developm.*, 1956, 27, 317-326.

DIFFERENCES BETWEEN MOTHERS' AND FATHERS' RATINGS OF LOW ANXIOUS AND HIGH ANXIOUS CHILDREN¹

KENNETH S. DAVIDSON, SEYMOUR B. SARASON, FREDERICK F.
LIGHTHALL, RICHARD R. WAITE, and IRVING SARNOFF
Yale University

As part of an ongoing project concerned with the development and correlates of children's anxiety, this study focuses on differences among parents' ratings of high anxious (HA) and low anxious (LA) children. Recent reports (1, 4, 5) from this project described the Test Anxiety (TA) and General Anxiety (GA) Scales for children² which were used to determine placement of children in the LA and HA groups in the present study. The previous reports have presented evidence pertaining to the validity of the anxiety scales. In the present study, validity of the scales was also of primary importance. That is, it was predicted that on items of a personality check list parents of LA children would rate their children more favorably than would parents of HA subjects. In addition, the present procedure permitted another important investigation. Most researchers in this area have relied on mothers' judgments of their children's characteristics, whereas both parents' evaluations of their child were obtained in this study. As a result, it was possible to compare the fathers' ratings with the mothers'. A further refinement was possible in that boys and girls were represented equally in the two anxiety groups. Thus, fathers' and mothers' ratings of LA and HA boys could be investigated as well as both parents' ratings for girls. While no specific differences were expected between fathers' and mothers' judgments, it is clear that any significant differences would have important implications for research relying on parental evaluations of children's personalities.

METHOD

Subjects

The children in this study were selected from 747 pupils in grades two through five who had answered the TA and GA questionnaires in the previous grade. Additionally, all had taken the Otis Alpha Intelligence

¹ The project of which this study is a part is being supported by a grant from the United States Public Health Service (M-712).

² Copies of these two scales are available for research purposes only.

CHILD DEVELOPMENT

Scale. Children were excluded whose parents were separated or divorced or who were regarded as arithmetic, reading, or behavior problems by their teachers. Subjects were considered as HA whose TA and GA scores were in the upper quartile of the distributions, and those whose scores fell in the lower quartiles were placed in the LA group. Thirty-two pairs of children were obtained by matching LA with HA subjects on grade, sex, and Otis Alpha T scores. Four such pairs were matched for each grade-sex group and the design was a 2 (anxiety) by 2 (sex) by 4 (grade) arrangement, with four subjects in each cell.

Procedure

The parents' ratings of their child were obtained by one of us (K.S.D.), who interviewed them in their homes at their convenience. In the first interview, the mother responded to questions about the child's school experiences, and the parents filled out an attitude scale on child rearing. A second interview with the mother followed about two months later and pertained to the child's experiences as well as personality in general. After the first item, a general free description of the child, the parent or parents filled out the personality check list.

This check list consisted of 25 items representing a variety of personality characteristics. For each item, a rating was required on a six-point scale with definitions for the extremes provided. For example, item 1 was presented as follows:

Immature	3	2	1	1	2	3	Mature
Acts younger than age							Acts grown up

Each parent was instructed to compare their child with children of his own age. They were to decide which of the two words described him better. Then they circled one of the three numbers on that side, with the three numbers defined as 1, a little more on this side; 2, definitely on this side; and 3, very much on this side. There were 16 items with clearly favorable and unfavorable extremes which were placed randomly in the list with the favorable extremes put on the right and left in random fashion.

If the father was not present, the check list was left for him with appropriate instructions and the explanation that the mother had expressed her own opinions and that she should not influence the father's rating of the child. In most cases the fathers were not present, but cooperation was excellent and the ratings by both parents of the 64 subjects were obtained.

RESULTS

All analyses of the evaluations were directional only and did not take account of the size of the difference between ratings. The procedure involved comparing the ratings of each matched pair of subjects on each of

the 16 items previously judged by three people as clearly having a favorable value at one extreme and an unfavorable value at the other. For the 16 pairs of HA and LA boys, the fathers' ratings of each matched pair were examined to see whether the father of LA rated his son more favorably or unfavorably than the father of the HA judged his boy. For each item, the number of times the father of the LA boy rated him more favorably than the HA boy's father judged his son (matched with the LA son) was counted as well as the number of times the LA boy was judged less favorably than his mate. The obtained proportion was subjected to a sign test and the p value was computed. This test was used for each of the 16 items. The instances in which the fathers of a matched pair of LA and HA boys rated their sons equally favorably were not considered in obtaining the

TABLE I

DIFFERENCE IN PARENTS' FAVORABLE RATINGS OF LA AND HA CHILDREN*

Item†	Father -Boy	Mother -Boy	Father -Girl	Mother -Girl	Parent -Boy	Parent -Girl	Father -Child	Mother -Child	Parent -Child	Mean Diff- erence Test‡
Mature	.02	.29	.29	.78	.02	.50	.02	.50	.07	.025
Sociable	.29	.50	.50	.50	.35	.50	.35	.50	.44	.05
Unanxious . . .	.28	.50	.30	.30	.42	.28	.16	.43	.18	.001
Optimistic . . .	.39	.39	.07	.50	.27	.15	.08	.42	.12	.001
Responsible . .	.07	.50	.09	.50	.15	.10	.02	.50	.04	.005
Active	.71	.50	.50	.75	.75	.76	.75	.76	.87	.995
Relaxed	.60	.71	.13	.29	.77	.09	.36	.50	.39	.20
Leads Others . .	.19	.06	.50	.50	.03	.50	.27	.20	.12	.005
Independent . .	.61	.50	.71	.50	.65	.58	.77	.50	.71	.995
Generous	.50	.81	.97	.72	.74	.98	.92	.90	.98	.999
Ambitious . . .	.11	.13	.50	.78	.03	.60	.13	.41	.14	.05
Affectionate . .	.87	.77	.83	.39	.92	.59	.95	.50	.88	.995
Adaptable . . .	.81	.60	.13	.19	.83	.06	.50	.43	.38	.25
Care in Plans . .	.28	.13	.39	.81	.14	.58	.29	.50	.27	.10
Not Sensitive . .	.40	.29	.15	.50	.22	.16	.28	.27	.09	.001
Takes Blame . .	.39	.61	.50	.50	.50	.42	.34	.50	.44	.15
Mean Differ- ence Test§	.01	.25	.15	.60	.20	.20	.10	.50	.10	.002¶

* Probabilities below .50 indicate that the LA were rated more favorably than the HA and values over .50 indicate that the HA were judged more favorably than the LA subjects. Items significant in the predicted direction at or below $p = .10$ are italicized as well as those items significant in the opposite direction with a p value of .90 or above.

† The terms for the favorable extremes of the items are listed.

‡ $N = 9$.

§ $N = 16$.

¶ $N = 144$.

probabilities, since each pair of LA and HA children was matched carefully on intelligence, sex, and grade. Following the same procedure, the ratings of the LA subjects were compared with those of the HA subjects for the various parent-child groupings given in the first row of Table 1. In all the groupings, mothers' ratings were compared with mothers', and fathers' compared with fathers'.

The results of these analyses are presented in Table 1 in the form of probabilities. For each parent-child grouping, there is a column of p values, one for each of the 16 items. At the bottom of each column, there is a significance point determined by a t test computed to evaluate the mean difference from chance of the 16 p values in each column. In addition, for each item, t tests were computed to determine the significance of the mean difference from chance of the nine probabilities obtained for the various breakdowns of parent-child groups (column 10). Finally, the mean difference test was done for all 144 p values in the main body of the table.

For all 144 p values, the mean difference from .50 is significant at almost the .002 level. The same results are reflected in the column of parents' ratings of children, where only four items yield results in the wrong direction while 12 are in the predicted direction, and the mean difference test is significant below the 10 per cent point of confidence. Of the nine mean difference tests for the columns, seven are in the predicted direction. Over-all, the evidence indicates that parents of LA children rate them more favorably than parents of HA children rate their children on these variables.

The results of the comparison of ratings of LA by their mothers with ratings of HA by their mothers were pooled with the results of the comparison of fathers' ratings of LA with ratings of the HA by their fathers. When this is done, the results tend to significance in the predicted direction (column 9). When the judgments of parents are pooled, their ratings of LA boys and of LA girls are more favorable, though not significantly, than parents' judgments of HA boys and HA girls (columns 5 and 6). However, in comparing fathers' with mothers' evaluations, significance is attained and approached only for the fathers. According to the mean difference test, the fathers of LA subjects rate them more favorably than the fathers of HA subjects rate them. That result is significant for the fathers-boys column and tends to significance for the fathers-children and fathers-girls categories. By that test, the mothers of LA do not judge them more favorably than mothers of HA evaluate their children beyond chance, and in fact, to a slight degree, mothers of HA girls rate them more favorably than their LA mates are judged by their mothers. In other words, HA and LA children tend to be discriminated in the predicted direction by fathers but not by mothers.

When attention is focused on individual items, there are three items on which fathers discriminate between HA and LA in the predicted direction, while mothers do not (columns 7 and 8). There is another item on which fathers also discriminate between HA and LA but not in the pre-

dicted direction. In short, whatever discriminations are made are the results of the fathers' ratings and not the mothers'.

In column 5 are given the results of a comparison of the ratings by parents of HA and LA boys. The corresponding results for girls are in column 6. When these two columns are viewed in relation to each other, two things seem worthy of note. First, there are certain items (mature, leads others, generous, affectionate) in which HA and LA boys are discriminated but girls are not, or HA and LA girls are discriminated but boys are not. Second, there are two items (relaxed and adaptable) in which LA girls are rated more favorably than HA girls, while on the same items HA boys are rated more favorably, though not significantly, than LA boys. There is also one item (ambitious) on which LA boys are rated more favorably than HA boys, whereas HA girls are rated slightly more favorably on that item than LA girls. Another way of stating the above is that there is not a single item which discriminates significantly between HA and LA boys as well as between HA and LA girls. This finding does not permit of any clear-cut conclusion. It does suggest, however, that the anxious boy as defined in this project is perceived by parents differently than the anxious girl.

DISCUSSION

The most suggestive finding reported above concerns the failure of mothers' ratings to discriminate between the anxiety groups whereas the fathers' ratings do discriminate in the predicted direction. This result takes on added significance in light of the fact that significant and predicted differences between these 32 matched pairs of children have been obtained in other situations: direct classroom observations, the Rorschach situation, and an experimental learning task³ (3, 4, 5).

Several explanations, of course, are possible. It may be that fathers, who spend less time with their children than do mothers, are less emotionally involved with their children and perceive them somewhat more dispassionately. It may be also that the mothers, who previous to the ratings had been asked many questions about the adjustment of their child, were in some way compensating for or reacting against whatever facts or attitudes they had revealed to the interviewer. It is also conceivable that the ratings of the mothers and fathers are not discrepant in that they may be using samples of behavior which differ in terms of time (e.g., day vs. night), content (e.g., observation of peer relationships) and variety of possible reactions. It is one thing to say that the ratings of mothers and fathers are different and another thing to say that one is valid and the other is not. The only conclusion we feel entitled to draw is that developmental studies which

³ In the study involving the experimental learning task, 24 of the 32 matched pairs of children were used for reasons cited in that publication.

rely on the ratings or observations by the mother, and this characterizes almost all such studies, may in overlooking the father be robbing themselves of important data.

It was noted earlier that while there were certain items which discriminated between HA and LA boys and others which discriminated between HA and LA girls, there was not a single item which discriminated for both sexes. We interpret this as suggesting that anxiety, as defined by scores on our two scales, has different psychological significance for boys and girls. The following very brief summary of findings of other of our studies may clarify this interpretation. Although girls consistently get higher scores than boys on both the Test Anxiety and General Anxiety Scales, the predicted interfering effects of anxiety tend to show up much more clearly between HA and LA boys than between HA and LA girls. We have elsewhere (4) suggested that our culture makes it easier for a girl to admit to anxiety than for a boy—one might say that we expect and support the admission of anxiety in a girl far more than in a boy. We tentatively hold to the opinion that boys who score high anxious are admitting to reactions which they experience as ego-alien to a greater degree than is true of girls. This opinion is not incompatible with the finding in the present study that no item of the rating scales discriminated both between HA and LA boys and between HA and LA girls.

REFERENCES

1. SARASON, S. B., DAVIDSON, K. S., LIDTHALL, F. F., & WAITE, R. R. A test anxiety scale for children. *Child Develpm.*, 1958, 29, 105-113.
2. SARASON, S. B., DAVIDSON, K. S., LIDTHALL, F. F., & WAITE, R. R. Direct classroom observations of high anxious and low anxious children. Unpublished data.
3. SARASON, S. B., DAVIDSON, K. S., LIDTHALL, F. F., & WAITE, R. R. Rorschach behavior and performance of high anxious and low anxious children. *Child Develpm.*, in press.
4. SARNOFF, I., LIDTHALL, F. F., WAITE, R. R., DAVIDSON, K. S., & SARASON, S. B. A cross-cultural study of anxiety amongst American and English school children. *J. educ. Psychol.*, in press.
5. WAITE, R. R., SARASON, S. B., LIDTHALL, F. F., & DAVIDSON, K. S. A study of anxiety and learning in children. *J. abnorm. soc. Psychol.*, in press.

CHILD DEVELOPMENT

Vol. 29
June, 1958
No. 2

Editor

WILLIAM E. MARTIN
Purdue University

Consulting Editors

BETTYE McDONALD CALDWELL
Washington University School of Medicine

ERCEL S. EPPRIGHT
Iowa State College

EUGENE S. GOLLIN
Queens College

HARRY LEVIN
Cornell University

PAUL H. MUSSEN
University of California

KATHERINE E. ROBERTS
Merrill-Palmer School

WATARU W. SUTOW
University of Texas

WILLIAM W. WATTENBERG
Wayne State University

Managing Editor

KATE HOFFMAN

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS
(of The Society for Research in Child Development, Inc.)
Purdue University
Lafayette, Indiana

Made in United States of America

Notice to Contributors

Child Development is devoted to basic research and theory concerning the growth and development of the child.

Manuscripts should be submitted in duplicate, each copy bearing only the title as a means of identification. The title, author's name, and institutional affiliation (together with any footnotes which reveal the identification of the author) should be submitted on separate pages.

Contributors are asked to refer to the *Publication Manual of the American Psychological Association* (1957 Revision) in the preparation of manuscripts.

The following rules require special emphasis:

1. All copy, including indented matter, should be typed double-spaced on white standard paper. Lines should not exceed six inches in length.

2. Footnotes should be few in number. They should be indicated in the text by superscript arabic numerals, numbered consecutively throughout the manuscript. The footnotes themselves should be typed double-spaced on a separate page placed at the end of the manuscript.

3. Tables and figures are costly to print and should be used only when essential. The author will be asked to contribute to the cost of composition of tabular and graphic material, which is excessive in amount or unduly complicated in form. Each table should be typed on a separate sheet. Insert a guide line (e.g., Table 1 about here) at the appropriate place in the manuscript. Figures should be prepared with black India ink, suitable for photography. Legends typewritten on an ordinary typewriter are unacceptable.

4. References in the text should be indicated by numbers in parentheses that refer to the list of references at the end of the article. Do not use footnotes for this purpose. In the list of references, the arrangement should be alphabetical by author's surname, or for multiple authors by the name of the senior; for any one author, by order of date. The following standard form should be observed:

Jones, H. E. Development in adolescence. New York: D. Appleton-Century, 1943.

Shepard, Winifred O., & Schaeffer, M. The effect of concept knowledge on discrimination learning. Child Developm., 1956, 27, 173-178.

The World List of abbreviations is used for journals.

The manuscript submitted should represent the final form in which the author wishes the paper to appear. All changes in proof (other than printer's errors) will be charged to the author.

All manuscripts should be addressed to William E. Martin, Editor, Child Development Publications, Purdue University, Lafayette, Indiana.

AGE AND SEX DIFFERENCES IN MOTIVATIONAL ORIENTATION TO THE COMMUNICATIVE ACT

RONALD PENNY

Australian National University

In our everyday communication relationships we constantly make inferences, perhaps not always consciously, about the communicator's motivation from both what he says and how he says it. We observe the fact that he repeatedly returns to particular topics, that he selects some aspects of a topic and disregards others, that he uses a particular tone of voice or style of speech, and so on. From such data we may infer that he has a "one track mind," that he is biassed, that he is being sincere or insincere. Probably most inferences of this kind are made automatically, on the basis of very limited data together with rather broad assumptions about human nature. We rarely go to the trouble of verifying our inferences by further observation or by seeking corroboration of their validity from other persons. One result of this imprecision is that we have little knowledge of developmental trends concerning the aim or orientation of communication. This paper is a report of an experiment directed toward that end. We shall present results which throw some light on the motivational orientation of statements made by Australian school children concerning the content of their communications on various topics.

PREVIOUS RESEARCH

A great deal of evidence concerning sex differences in content was accumulated during the 1920's and 1930's (1, 4, 5, 7, 9). In the main, the data were collected by recording snatches of conversation overheard in the street. Attempts were made to determine the effects of various independent variables such as the communicant's social status, the sex-composition of the communicating pair (man-to-man, man-to-woman, etc.), and the social setting within which the communication took place, but the dependent variable was usually simply the amount of "thing-related" and "person-related" content. On the whole, the results of these studies uniformly indicate the presence of sex differences; men tend to talk about "things" and women about "persons." Many other studies in related fields such as interests, values, reading matter, and movie preferences are congruent with those concerned with conversation content.

The findings of Piaget (8) in connection with the language behavior of children and those of other investigators who have applied his methods of categorizing language behavior (2, 3, 6) are also relevant to our present

concern, for here we have some implicit emphasis on the functions served by various types of communicative acts. Language behavior is "egocentric" when the child "does not bother to know to whom he is speaking nor whether he is being listened to . . . (he) speaks only about himself . . . he feels no desire to influence his hearer nor to tell him anything" (8, pp. 9-10). His speech is "socialized" when there is "interchange of ideas," criticism of the work or behavior of others, when there is definite interaction. Clearly, there are motivational implications to these definitions. In their present form, however, they are far from adequate for the purpose of making clear-cut categorizations of communicative acts from the point of view of their motivational orientations.

In this paper we shall be concerned with data which bear somewhat more directly on the extent to which communications betray motivational orientations and which are related also to the earlier research concerned with sex differences in communication content.

METHOD OF CATEGORIZING MOTIVATIONAL ORIENTATION

By motivational orientation we mean something like this: we assume that all communicative acts are motivated and that the performance of the act partially or wholly satisfies the need. The question is, just what kind of need are we talking about? Alternatively, what kind of satisfaction is the communication aimed at? One way of answering this question is

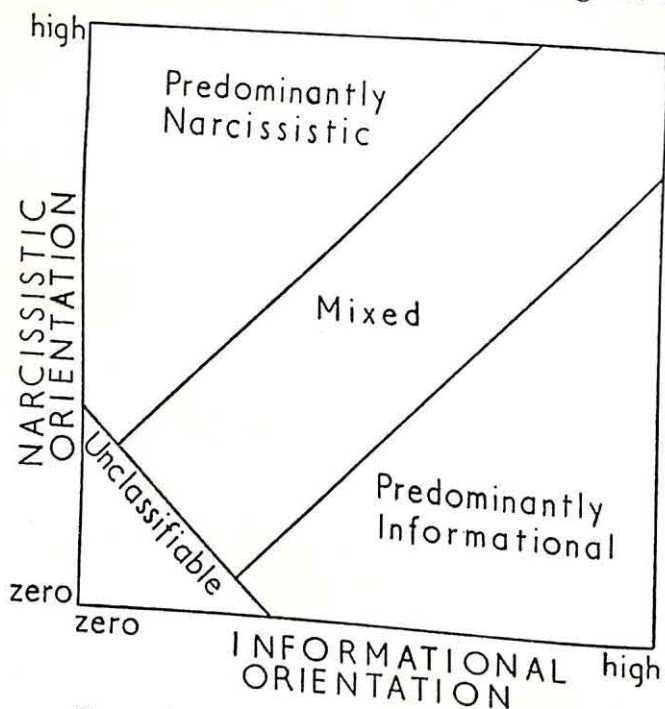


FIGURE 1—Motivational orientation categories.

to state the extent to which the communication is oriented toward information transfer and the extent to which it is concerned with self-involvement. We shall refer to these two dimensions as informational and narcissistic orientations. The reason for choosing these two dimensions rather than others is that they permit the locating of the intent of communications in a space constructed on the basis of two variables which are relevant to virtually all communications; furthermore, no communication will have a location of zero on both of these two dimensions. Thus, if we represent this space by Figure 1, every communication can be located at a point other than the zero-zero point. Clearly, we are dealing with only one way of describing motivational orientation. There are, of course, other dimensions that could be used in specific instances. For example, a communication such as "How about coming around to my place to play after school's out?" might indicate the presence of an affiliation motive but such specific dimensions would be relevant to comparatively few communications.

The problem of assigning precise informational and narcissistic values to communications is a difficult one. In the present analysis we have sidestepped this issue and have used a classificatory scheme indicated by the dividing lines in Figure 1. The definitions of the categories are:

Informational orientation—the communication content is predominantly concerned with the objective qualities of physical objects or persons. The inclusion of the self in the form of reference to one's own possessions or opinions must be slight compared with the objective "loading" of the communication.

Narcissistic orientation—the content of the communication is predominantly concerned with drawing attention to the self.

The "mixed" and "unclassifiable" categories will be referred to later.

TABLE I
CLASS SIZE AND MEAN AGE

	FIRST YEAR		SECOND YEAR		THIRD YEAR		FOURTH YEAR	
	N	Mean Age	N	Mean Age	N	Mean Age	N	Mean Age
Boys . . .	25	13-4	37	14-5	29	15-5	21	16-2
Girls . . .	28	12-11	19	13-11	20	14-11	16	15-8

SUBJECTS

The Ss comprised the first four grades of a boys' and a girls' grammar school, 112 boys and 83 girls. We shall refer to these grades as "First Year," "Second Year," etc. There was one class for each grade at each school. Table I shows the sizes of the classes and the mean ages of the members. The scholastic level of First Year corresponds to first year High School (Australian system).

CHILD DEVELOPMENT

DATA

The data were collected in individual interview sessions. Previous to the interview each *S* had filled in questionnaires in class concerning his communicatee selections for 12 different topics. These questions were in the form: "Whom do you talk to about . . . ?" The topics were: favorite hobby, what you want to do when you leave school, personal secrets, school work, dreams, father and mother, films and film stars, teachers, girls (for boys) or boys (for girls), jokes, school sport, and religion. In the interview the *S* was asked, with respect to 10 of these topics (jokes and personal secrets being excluded): "When you talk about . . . just what are the kinds of things you do talk about?" The responses, which constitute the raw data of this analysis, were recorded verbatim.

ANALYSIS

Every response was classified as being either predominantly informationally or narcissistically oriented with the exception of those that clearly contained both orientations ("mixed") and those that contained neither ("unclassifiable"). The following are examples for each of the categories.

Topic	Response
	<i>Informational Orientation</i>
Girls	"Where they come from, where they live, if they can dance, if they go out much, if they've got any brothers here, what sort of a girl she is—big, tall or thin."
School work	"Lessons, interest of the work, wheat seeds, contour-ploughing, oxy-welding."
Films and film stars	"The roles they're suited for, the clothes they wear and the places they go when they're not acting, all their husbands, sort of houses they live in, the way they entertain."
	<i>Narcissistic Orientation</i>
School sport	"Why I wasn't picked, that I always go to practice and never get picked."
Dreams	"What I dreamt about, adventures, girls, driving Dad's car and smashing it up."
Father and mother	"How Mum cooks, about going to work with my father, how I like being at home with them during the holidays, how much pocket-money they let me have."
	<i>Mixed Orientation</i>
Teachers	"If they give you the cane, if they are cranky, if you've got a good house-master, what they give you in class, when they take you next."
Boys	"The things they do like climbing trees, about them teaching us to do it, what school they go to, where we have seen them before."
	<i>Unclassifiable</i>
	"That's a secret."
	"I'm afraid I'm not very much help to you."
	"I used to talk about it but I don't now."

This kind of analysis can never be thoroughly objective because of the fact that the categorizations are not made on the basis of the structural aspects of the responses. In the absence of such objective classificatory criteria the projective component in the interpretations is likely to be of considerable importance. For this reason a reliability check was made by having a randomly selected sample of 50 responses from the boys' data and 50 from the girls' classified independently by an experienced anthropologist. The result of this check was as follows: 65 per cent complete agreement, 27 per cent partial disagreement (as when one judge classified a response as being informational or narcissistic and the other as mixed), 8 per cent complete disagreement (as when one judge classified the response as informational and the other as narcissistic). No significant judge-bias in the informational or narcissistic direction was present in the 27 partial disagreement instances. Whilst 65 per cent complete agreement is not high it was felt that this, together with the mere 8 per cent complete disagreement, was sufficient to justify the belief that the categorizations were not unduly biased.

RESULTS

Tables 2 and 3 show the percentages of responses in each of the categories. In Table 2 the data have been grouped in terms of age level. Such a grouping is not identical with a grouping in terms of class level (First Year, Second Year, etc.) as there was a considerable amount of age overlap between classes. Amongst the Second Year boys there were, for instance, 13 Ss between 13-0 and 13-11 years, 16 between 14-0 and 14-11, 7 between 15-0 and 15-11, and 1 aged 16-1. Also, as may be seen from Table 1, there was a six-months difference between the mean ages of the girls' classes compared with the boys' at each of the class levels. Table 2

TABLE 2
PERCENTAGES OF RESPONSES IN MOTIVATIONAL CATEGORIES
(Data grouped according to chronological age)

Age Level	P E R C E N T A G E S *									
	N		Informational		Narcissistic		Mixed		Unclassifiable	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
12-0 to 12-11 ..	6	18	53	31	13	24	15	26	18	19
13-0 to 13-11 ..	27	19	46	39	13	20	28	31	13	9
14-0 to 14-11 ..	31	21	42	43	15	17	26	29	17	11
15-0 to 15-11 ..	32	18	52	47	10	12	25	35	13	6
16-0 to 16-11 ..	16	7	55	43	8	10	24	41	14	6
17-0 to 17-11 ..	5	—	58	—	8	—	16	—	17	—

* Significance of sex differences (sign test): Informational, $p = .19$; Narcissistic, $p = .03$; Mixed, $p = .03$.

holds chronological age constant and is, therefore, a more accurate representation of sex differences in motivational orientation than Table 3 in which the data are grouped in terms of class level. Table 3 shows essentially the same trends as Table 2, though as has been pointed out, it is not really suitable for intersex comparisons. On the other hand, there are grounds for preferring it to Table 2 for intrasex age comparisons. The chief reason is that about 90 per cent of the Ss' communicatees were drawn from their school classes rather than their age groups. The answer to the question: "When you talk about . . . just what are the kinds of things you do talk about?" referred, therefore, to communication within the class group. Furthermore, there were, presumably, mental age reasons for the age overlap between classes. A 14-year-old boy in Fourth Year is presumably rather brighter than a 14-year-old in First Year. It seems quite plausible, therefore, that a grouping of the data in terms of class level is not only justifiable but also psychologically more meaningful for intrasex age comparisons than a chronological age grouping.¹

TABLE 3
PERCENTAGES OF RESPONSES IN MOTIVATIONAL CATEGORIES
(Data grouped according to class level)

P E R C E N T A G E S											
		N		Informational		Narcissistic		Mixed		Unclassifiable	
		Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
1st Year		25	28	42	31	18	24	26	28	14	17
2nd Year		37	19	46	42	13	17	26	31	15	10
3rd Year		29	20	56	46	10	12	22	32	12	10
4th Year		21	16	48	47	8	14	22	36	22	3

Differences Between Class Levels (Table 3)

1. The percentages in Table 3 indicate fairly conclusively that motivational orientation, as we have defined it, changes with increasing (mental) age. With only two minor exceptions (Fourth Year boys, informational, and Fourth Year girls, narcissistic) the informational percentages increase and the narcissistic decrease.

2. Within each sex group the informational percentages are greater than the narcissistic at all class levels. The differences, of course, become greater with increasing age.

¹ The total N is less in Table 3 than in Table 2 because of the fact that the members of the small Fifth Year classes have been included in Table 2 but, because of their sizes, not included in Table 3.

3. The mixed percentages for the boys decrease slightly with increasing age; those for the girls increase.

The same trends occur in Table 2 though less unequivocally.

Sex Differences (Table 2)

4. At each age level, with one exception (14-0 to 14-11), the boys tend to be somewhat more informationally inclined than the girls. In all cases the girls are more narcissistically inclined than the boys.

5. The mixed percentages for the girls are greater than those for the boys at all age levels. This would seem to indicate that for the girls there is a counteracting trend to that shown by the informational percentages. The degree of predominantly narcissistic orientation decreases with increasing age but instead of this orientation changing clearly to an informational one, as seems to be the case with the boys, more and more responses have to be classified as being of mixed orientation—perhaps a kind of narcissistic “contamination” of what would otherwise be informationally oriented responses.

DISCUSSION

Analysis of the interpersonal communication process ought to be especially useful from the point of view of throwing light on many aspects of the socialization process. This is because the performance of the communicative act involves a variety of decisions. Decisions have to be made in connection with the selection of the communicatee, the topic, the appropriateness of the social setting for the particular act, and so on. This is not to say that there is conscious deliberation before or introspective analysis after each act. The performance of the act does, however, clearly involve the interplay of many social and individual forces. One such force is, of course, the communicator's motivation.

In a sense, the communicative act is an experiment in which the communicator gets more or less immediate feedback concerning others' evaluations of him and his communication procedures. Such feedback must constitute an important aspect of the socialization process.

Motivational orientation, we have assumed, is betrayed by the mode of expression of the content of the communication. Presumably, this information is available to and reacted to by the communicatee thus providing the communicator with feedback concerning the acceptability of his motivational orientation. In this way motivational orientations may well be modified by the social forces existing within the peer group. There may be and probably are other modifying agencies. Parents and teachers may well react to what they perceive to be particular orientations in approving or disapproving ways. They may, for example, punish boastfulness.

Our data show that certain modifications to motivational orientation are taking place between the ages of 12 and 17. We are suggesting that these

modifications are being brought about, at least in part, by peer group forces. The communication data we have analysed are, after all, *peer group* communication data. We are not here concerned with the basis of these forces but rather with showing that they apparently exist and that they affect the sexes differently.

The general direction of the change is, as we have seen, from narcissistic toward informational. Although extrapolation on the basis of such data involves the assumption of certain developmental constants, it would seem to be a reasonable guess that for age groups lower than our First Year level the proportion of narcissistic responses would be greater. Piaget's findings support this guess. His "ego-centrism" is not synonymous with our narcissistic orientation category but a rearrangement of his data in such a way as to fit fairly well with our definitions indicates a narcissistic proportion of between 0.55 and 0.66 for children between the ages of seven and eight.² Similarly, we should expect that there would be further increases in the informational proportion with further increases in age. Both these tendencies are, of course, limited. No communication can be "pure" in the sense of lacking either motivational orientation completely. But that is not a psychological issue.

SUMMARY

It has been assumed that the motivational orientation of the communicative act is betrayed by the mode of expression of the content. A classificatory scheme and mode of analysis have been described and applied to content data obtained from school children of both sexes between the ages of 12 and 17. These data reveal certain age trends and sex differences. The relation of such an analysis to the investigation of the socialization process has been discussed.

REFERENCES

1. CARLSON, J. S., COOK, S. W., & STROMBERG, E. L. Sex differences in conversation. *J. appl. Psychol.*, 1936, 20.
2. FISHER, M. S. Language patterns of the pre-school child. *Child Develpm. Monogr.*, 1934, No. 15.
3. HENLE, M., & HUBBELL, M. B. "Egocentricity" in adult conversation. *J. soc. Psychol.*, 1938, 9.

² This rearrangement has been applied to the table which summarizes the data for the children Pie and Lev (8, Ch. 1, sec. 10) and consists of allocating collective monologue, criticism, and commands to our narcissistic category and adapted information and requests to our informational category. Because they cannot be unambiguously classified the following Piaget categories have had to be excluded: repetition, monologue, and answers.

4. LANDIS, C. National differences in conversations. *J. abnorm. soc. Psychol.*, 1927, 21.
5. LANDIS, M. H., & BURTT, H. E. A study of conversations. *J. comp. Psychol.*, 1924, 4.
6. MABIE, E. A study of conversation of first grade pupils during free play periods. *J. educ. Res.*, 1931, 24.
7. MOORE, H. T. Further data concerning sex differences. *J. abnorm. soc. Psychol.*, 1922, 17.
8. PIAGET, J. *The language and thought of the child*. London: Routledge and Kegan Paul, 1948.
9. STOKE, S. M., & WEST, E. D. Sex differences in conversational interests. *J. soc. Psychol.*, 1931, 2.

PREDICTION OF SOCIAL ACCEPTANCE IN COMMUNITY YOUTH GROUPS

HELEN R. MARSHALL

School of Home Economics, University of Kentucky

Community groups or clubs are a prominent feature of American life but have been ignored for the most part in research on social acceptance. No examples have been found by the author of their use in investigations of factors associated with social acceptance, although school, camp, industrial, and military groups have been frequently used for such studies. The research literature on these populations gives consistent support to some hypotheses concerned with factors affecting social acceptance. Enough variables appear to suggest that some be tested concurrently on a population different from any used originally. This initial step in (or test of) development of a theory of social acceptance is possible at the present time by use of the unexplored populations of community groups.

This investigation proposes to test the prediction for community youth groups of six hypotheses concerned with factors affecting sociometric acceptance. Specifically, the community youth group populations are as follows: (a) 23 Home Economics 4-H Clubs of 324 girls aged 10 to 20 years located in one Illinois county; (b) 22 Agriculture 4-H Clubs of 314 boys and 91 girls aged 10 to 20 years in the same county.

The six hypotheses were developed inductively from empirical evidence as research conclusions. They are listed below with a description of the measure or classification used in this investigation to define the factor. Complete citation of research literature is not attempted here; rather, reviews and reports are cited that list and describe the evidence for these hypotheses.

1. *In child or youth groups, age has a positive, low correlation (.15-.20) with sociometric status* (21). Within the school classes studied, standard deviations averaging around four months of age have been reported. Age had a possible nine-year range within the clubs of this investigation. Hence, age difference in months from the median age of the club is used as a measure of the factor of age in addition to chronological age in months.

2. *Sociometric status in both sex groups is higher for girls than for boys.* Sociometric scores are used in the sex comparisons of this investigation and have been so used earlier in only one report (10). It has been frequently reported that boys choose girls more often than girls choose boys on social acceptance tests (2, 5, 6).

3. *Level of performance or achievement in the skills and knowledge essential for or being taught to the group is positively related to sociometric status* (15). Two performance ratings are included in the 4-H Club program. One, the project rating, is made at the county contest by trained

judges of an article (clothing, food, home furnishings) made by the girl as a part of project requirements in home economics clubs, or of an animal, fowl, or produce raised by agriculture club members. The home economics ratings of A, B, and C are comparable to grades of skill and knowledge used with other groups. Home economics project requirements attempt to adapt to age differences in performance. The five place agriculture ratings probably are not performance measures; i.e., quality of animals is affected by heredity so that ratings of these animals are not expected to relate to sociometric scores of club members. 4-H members may enroll in several projects, so any project rating of A placed the member in the A classification in this investigation.

The other performance rating is an end-of-the-year (September) project honor evaluation of members' ratings in all competitions and participation in all club activities and in community and school activities. Members must record activities on the project honor form to be considered, so that a member's submission of a filled out form does not necessarily imply performance in the rated aspects superior to all members not submitting these forms. Score points given to each activity and rating have been determined in past years by state extension specialists for use throughout the state. The forms submitted were scored by club leaders and committees of the county sponsoring groups at a county meeting. Scores of 90 to 100 were classed as State Project Honor, and of 75 to 89 as County Project Honor in home economics clubs. Classification was similar in agriculture clubs.

4. *Leadership is positively related to sociometric status (15).* This relationship has been reported for such definitions of leadership as near-sociometric choices by peers and ratings by superiors or observers of group interaction. Election to one of the five club offices is the definition of leadership used in this investigation. This definition is presumed to be comparable to peer leader nominations on near-sociometric tests. The author was unable to find any reports of the sociometric status of elected leaders in the groups of their election other than her report on women's clubs in this county (13). The second performance rating used to test hypothesis 3, the observers used in many investigations of leadership. Possibly the test of this hypothesis would be more comparable to the supporting evidence if the project honor classification were used to define leadership rather than the less ambiguous election to office dichotomy.

5. *Family and community variables may relate to social acceptance in particular groups.* Studies report conflicting evidence on the presence and absence of relations between social acceptance and classifications of socioeconomic status (1, 3, 11, 16, 20), sibling status or number (1, 4, 9, 13, 21), place of residence (17), and race or nationality (7, 18, 19). Two dichotomous classifications of this type are used in this investigation: residence in town or on a farm, and attendance at the same or different schools.

6. *Sociometric status in one group has a positive and fairly high correlation (.50-.60) with sociometric status in groups of different membership* (14, 22). Sociometric scores were obtained in both agriculture and home economics clubs for 33 girls, and within the county group attending District 4-H Club Camp for 15 members with scores in one of the two club populations.

SUBJECTS AND PROCEDURES

All individuals and groups used as subjects (Ss) resided in one Illinois county. In the 1950 census, 87.5 per cent of the 25,000 county population was classified as rural, a city of about 3,000 population being classified as urban. The county was close to the rural Illinois average in home facilities, family income, value of farm products sold, and percentage of farm tenancy, and may be described as having the narrow range of socioeconomic status characteristic of rural areas in this section of the United States. It ranked in the upper fourth of Illinois counties in such measures of level of education as that 9.4 per cent of the farm adults aged 25 years or more had attended college, and a median of 9.1 school years had been completed by these adults.

The groups used in this investigation were 4-H Clubs, the youth organization for education in agriculture and home economics of the Extension Service of the University of Illinois and the U.S. Department of Agriculture. 4-H Clubs have been organized in this county for more than 25 years. Membership in each club tends to be limited to a specific geographical area, such as a township. Each township in the county had either an agriculture or home economics club, and the majority had both types of clubs.

The Home Economics 4-H Club members serving as Ss in this investigation were the 270 girls present at the regular club meeting at which the author administered the sociometric test. Tests were administered to all 23 clubs within a six-weeks period of June and early July, 1956, prior to all competitions. Ss constituted 83 per cent of the enrollment. Club size ranged from five to 28 members and Ss in each club from four to 27 girls. The mean number of meetings held by the club at the time of test administration was 7.26 with an *SD* of 1.86, and a range of three to 10. The age distribution of home economics Ss is presented in Figure 1.

The Agriculture 4-H Club members serving as Ss in this investigation were the 87 boys and 47 girls present at the regular meeting of eight clubs at which the author or the assistant farm or home advisers (in three clubs) administered the sociometric test. Selection of eight of the 22 clubs was biased by the attempt to include clubs in which home economics Ss were enrolled, but otherwise was determined by the possibility of test administration at regular meetings within six weeks of the test administration to the Home Economics 4-H Club in that township. Agriculture Ss constituted 81.5 per cent of the enrollment of the eight clubs. Girl Ss were 52 per cent of the girls enrolled in all agriculture clubs and boy Ss were 28

CHILD DEVELOPMENT

per cent of the total enrollment of boys. Size of the eight clubs ranged from 13 to 40 members, and Ss in each club from 12 to 38. The mean number of meetings held by the clubs at the time of test administration was 5.88 with an SD of 1.46 and a range of four to nine. The age and sex distribution of agriculture Ss is presented in Figure 2.

Thirty-one members of the two club populations attended District 4-H Club Camp in mid-July for one week. The assistant home adviser individually administered the sociometric test to 26 members of the county group on the last day of camp attendance. Club scores were available for 15 of these campers, 13 girls from nine home economics clubs and two boys from the same agriculture club. Three girl campers were included among the 33 girls with scores in both agriculture and home economics clubs.

The measure of social acceptance was a three-question and three-choice form of the sociometric test. Each S was asked to choose three members present in his (her) group as answer to each of three questions: (a) With whom would you like to serve on a committee (any kind)? (b) With whom would you like to go to a county meeting, such as the spring rally, or the county show, or a demonstration school? (c) Whom do you like to sit with or talk to at your club meeting? Instructions stated that the same or different members could be named for each of the three questions.

In scoring the test, first choice for any question was given three points, second choice two points, the third choice one point. The sociometric score was one plus the sum of the weighted choices of the S by other Ss of the group.

The mean sociometric score for any group was necessarily 19 points since each S gave 18 points to others. Hence, Bartlett's test of homogeneity of variance was used to determine whether the scores in the clubs of each population were comparable. Differences in score variance of the 23 Home Economics 4-H Clubs and of the eight Agriculture 4-H Clubs were no greater than could have been expected by chance. It was possible, therefore, to treat sociometric scores of all Ss from one population as a single sample in analyses of relations with other factors. The population SD for sociometric scores was 11.2 in home economics clubs and 11.8 in agriculture clubs.

All other data were obtained from 4-H Club records kept in the county extension offices. Information used had been checked for accuracy by at least two of the adults most informed on that aspect: parents, leaders, or extension office personnel.

RESULTS AND DISCUSSION

Hypothesis 1. In child or youth groups, age has a positive, low correlation (.15-.20) with sociometric status.

This hypothesis predicted the correlations obtained between sociometric scores and CA in months for these community youth groups, as shown in Table 1. The mean and median sociometric scores for each year of

HELEN R. MARSHALL

age (10-0 to 10-11, etc.) are plotted for home economics clubs in Figure 1, and, with sex separation, for agriculture clubs in Figure 2. Hypotheses that the relationship was linear could not be rejected for CA or the age deviation score (to be described) in either club population.

TABLE I

PRODUCT-MOMENT CORRELATIONS OF AGE WITH SOCIOMETRIC SCORES

Age Measures	SOCIOMETRIC SCORES			
	Home Economics 4-H Clubs N Girls = 270	Agriculture Boys N = 87	4-H Girls N = 47	Clubs Total N = 134
Chronological age in months	.15*	.18	.42**	.22**
Months deviation + 50 from median age of club	.29**	.29**	.24	.22**

* r is significant at less than the .05 level.

** r is significant at less than the .01 level.

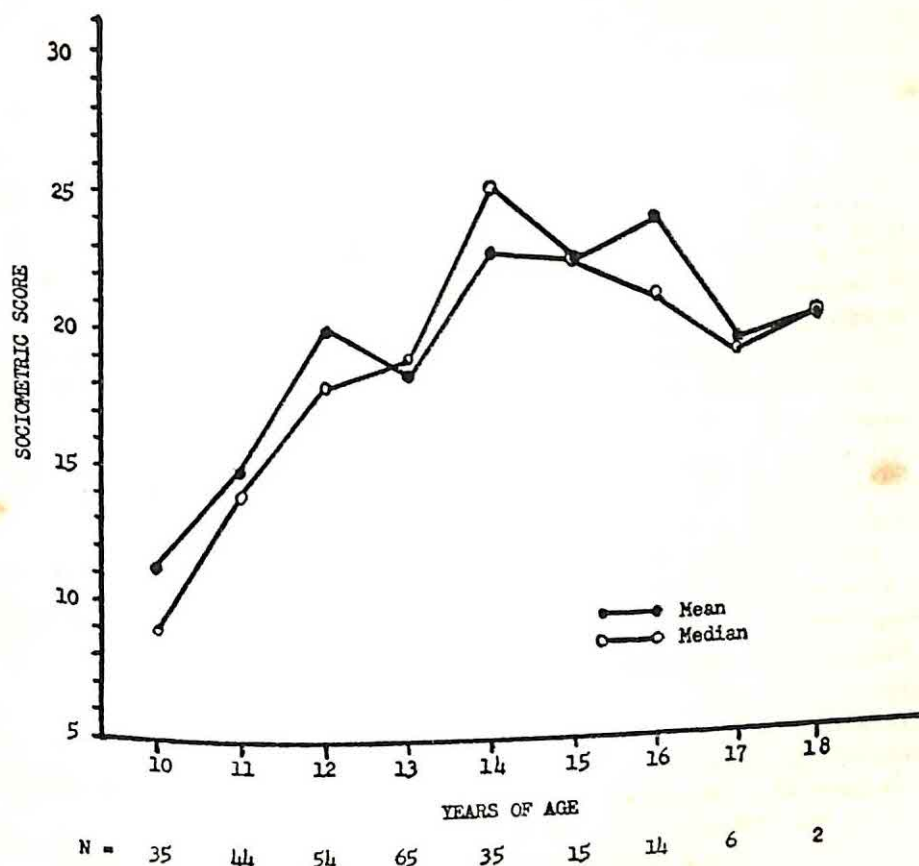


FIGURE 1—Relation of sociometric scores of Home Economics Club Ss to chronological age.

CHILD DEVELOPMENT

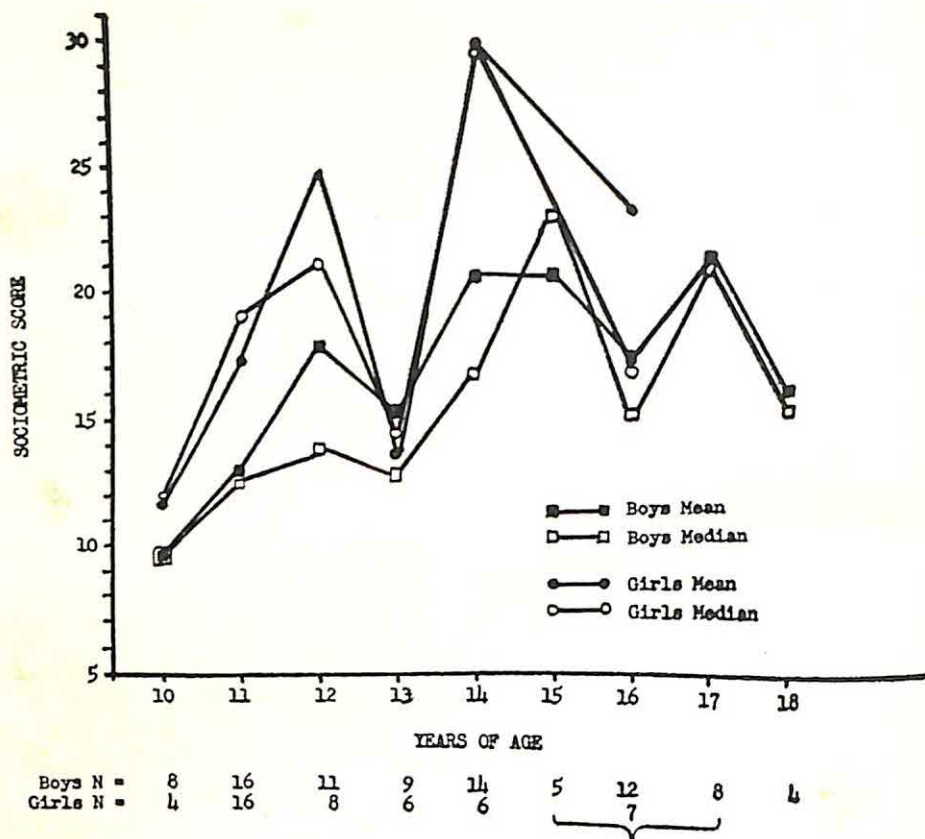


FIGURE 2—Relation of sociometric scores of Agriculture Club Ss to chronological age.

The mean age of home economics Ss was 13.3 years with an *SD* of 15.8 months, and the median age was 13.0 years. Within the 23 clubs, the median age ranged from 11.6 to 15.2 years, and the age range within the club was at least six years in all but four clubs. These descriptive statistics suggest that age deviation in months from the median age of the club might be a more useful measure of the relation of age to sociometric scores than *CA per se*. The measure used was months of deviation plus 50, to eliminate negative scores.

The correlation of months of deviation + 50 with sociometric scores was .29, as shown in Table 1, and the relationship is plotted in Figure 3. An almost straight line connects the mean sociometric scores below the median age and is reflected in the coefficient of .48 obtained between the two measures for the 88 Ss seven months or more younger than the median age of the club. This coefficient differs significantly from the *r* of .07 obtained for the 91 Ss with no more than six months difference in age from the median of the club, and from the *r* of .04 obtained for the 91 Ss seven months or more older than the median age of the club. The regression

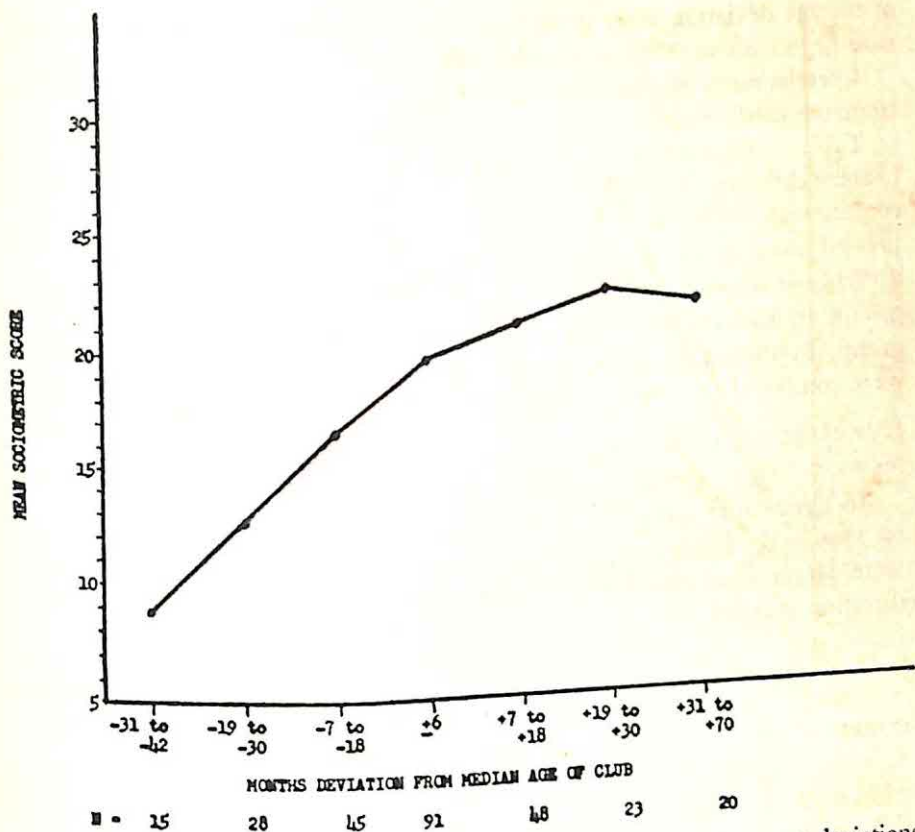


FIGURE 3—Relation of sociometric scores of Home Economics Ss to age deviations from the median age of the club.

equation fitted by the least squares method for predicting sociometric score (Y) from the minus seven or more months of deviation + 50 score (X) is: $Y = .989X + (-16.66)$.

The null hypothesis was rejected at the .001 level for the mean sociometric scores of these three groups in an F test of a simple randomized analysis of variance (8). The null hypothesis was rejected at the .02 level for the difference in mean sociometric scores of the below median age Ss and the median age Ss, and at the .001 level for the difference in means of below median age Ss and above median age Ss. The null hypothesis could not be rejected for the difference in means of the median age Ss and the above median age Ss.

It is concluded that in these girls' clubs deviations in age below the median age of the club are inversely related to sociometric scores, but deviations in age above the median age of the club do not relate to sociometric scores.

To obtain sociometric scores that were independent of age influences for use with other factors in this investigation, each home economics sociometric score was expressed as a deviation from the mean sociometric score

of the age deviation array groups presented in Figure 3 (12), and, to eliminate negative scores, 50 was added to each deviation.

Correlation coefficients of sociometric scores with months of deviation from the median age of the club + 50 in agriculture clubs are presented in Table 1. Because both age measures correlated with sociometric scores in these clubs and sex differences were present (to be described), a statistical control was used for both age and sex. Each sociometric score was expressed as a deviation from the mean sociometric score of the age-sex array groups presented in Figure 2 + 50, with three exceptions: the groups in the 13- and 14-year arrays were combined for girls, and the groups in the 15- and 16-year arrays, and in the 17- and 18-year arrays were combined for boys.

Hypothesis 2. Sociometric status in both sex groups is higher for girls than for boys.

In agriculture clubs, the mean sociometric score of girls was 20.1, $SD = 11.0$, and of boys was 16.7, $SD = 10.3$. The girls' sociometric scores were larger than the boys'; the null hypothesis of a difference in this direction was rejected at the .05 level. The difference in sociometric scores

TABLE 2
CORRELATIONS OF SOCIOMETRIC SCORES IN HOME ECONOMICS 4-H CLUBS
WITH DICHOTOMIES OF OTHER FACTORS
 $N = 270$

Dichotomy of Factor	Correlation Index	Sociometric Score	Sociometric Score Corrected for Age Differences
Project rating A (65%)—B or C (32%) (No exhibit—3%)	r_{bis}	.43**	.31**
County Project Honor Yes (38%)—No (62%)	r_{bis}	.24**	.26**
State Project Honor Yes (20%)—No (80%)	r_{bis}	.17**	.26**
Elected an officer of club Yes (40%)—No (60%)	$r_{pt\ bis}$	.17**	.12
Farm or town residence (in 13 clubs with both) N farm = 118 (63%) N town = 68 (37%)	$r_{pt\ bis}$	.04	.03
Difference in school (19 clubs had majority from one school) N same = 168 (74%) N different = 60 (26%)	$r_{pt\ bis}$	.08	.02

* r is significant at less than the .05 level.

** r is significant at less than the .01 level.

occurred in spite of the fact that the girls were younger than the boys and age had a positive relation with sociometric scores for both sexes. The mean age of girls was 12.8 years and the mean age of boys was 13.11 years. The null hypothesis of an age difference in this direction was rejected at the .001 level.

The sex difference in sociometric scores in community youth groups agrees with the hypothesis. The hypothesis specifies sociometric status rather than social acceptance. This distinction is made because an investigation with preschool children (10) found the null hypothesis could not be rejected for sex differences in observed social acceptance, observed social participation, and teacher judgments of social acceptance when the sex difference in sociometric scores was large and significant.

Hypothesis 3. Level of performance or achievement in the skills and knowledge essential for or being taught to the group is positively related to sociometric status.

Findings with the 4-H Clubs agree with this hypothesis, as shown in Tables 2 and 3. The home economics project rating correlated beyond

TABLE 3
CORRELATIONS OF SOCIOMETRIC SCORES IN AGRICULTURE CLUBS
WITH CLASSES OF OTHER FACTORS

Classes of Factors	Correlation Index	SOCIOMETRIC SCORES			
		Boys N = 87	Girls N = 47	Total N = 134	Corrected for Age and Sex, N = 134
Project rating (of 103 Ss exhibiting project) A (50%)—B (37.5%)— C or less (12.5%)	r_{tris}	.03	.13	.01	.04
County Project Honor Yes (18%)—No (82%)	$r_{\text{pt bis}}$	.25*	.28	.28**	.23**
State Project Honor Yes (8%)—No (92%)	$r_{\text{pt bis}}$	.30**	.31*	.30**	.26**
Elected an officer of club ... Yes (22%)—No (78%)	$r_{\text{pt bis}}$	.31**	.21	.27**	.16
Farm or town residence ... (in five clubs with both) N farm = 80 (84%) N town = 14 (16%)	$r_{\text{pt bis}}$	..	..	.14	.11
Difference in school (four clubs had majority from one school) N same = 49 (63%) N different = 29 (37%)	$r_{\text{pt bis}}$	.04	.10	.08	.07

* r is significant at less than the .05 level.

** r is significant at less than the .01 level.

chance with sociometric status while the similar agriculture rating, described as a non-performance measure, did not do so. In both club samples the county and state project honor evaluations had significant correlations with sociometric scores. Knowledge of the judges' rating or evaluation could not have affected these relations since sociometric tests were administered prior to these county competitions.

Hypothesis 4. Leadership is positively related to sociometric status.

This hypothesis did not predict correlations between election to club office and sociometric scores when the effects of age were controlled in these community youth groups, as shown in Tables 2 and 3. Significant correlations were obtained only when the positive relation of age to sociometric scores (and possibly election to office) was uncontrolled. This finding cannot be attributed to the particular county; a relationship was found to exist for the elected leaders in women's clubs in the same communities and county (13). It is possible that election to club office has a different meaning (relation to other variables) in developmental years than in adulthood. Possibly, also, peer nominations of leaders on near-sociometric tests may not resemble elections of leaders. If leadership had been defined as the for these groups. As stated in the description of measures, the project honor definition is comparable to definitions used in many investigations reporting this hypothesis as a research conclusion.

Hypothesis 5. Family and community variables may relate to social acceptance in particular groups.

The dichotomies of farm-town residence and of attendance at the same or different schools were not associated with social acceptance in these community youth groups, as shown by correlations in Tables 2 and 3.

Hypothesis 6. Sociometric status in one group has a positive and fairly high correlation (.50-.60) with sociometric status in groups of different membership.

The 33 Ss who were members of both types of clubs had a below average mean score, 16.1, $SD = 8.8$, in home economics clubs and an above average mean score, 20.1, $SD = 12.5$, in agriculture clubs. The respective age and age-sex corrected mean scores were 46.9 and 50.6. The hypothesis of no difference in these means was rejected at the .05 level in a one-tailed test for both the original and corrected scores. The product-moment correlation coefficient was .62 between the two sociometric scores, and was .49 between the sociometric scores corrected for age or for age-sex differences.

The correlation coefficients between sociometric scores in the county camp group and club were .55 for both the 13 girls and all 15 Ss. When the club score corrected for age or age-sex differences was correlated with the camp sociometric score, coefficients were .38 for the 13 girls and .42

for all 15 Ss. The reduction in size of the latter two coefficients suggests that age differences probably were associated with camp group sociometric scores also. Campers' ages were restricted to 10 to 14 years. Club sociometric score means were 17.8 and 18.5 for the 13 and 15 Ss respectively, and camp sociometric score means were 18.3 and 18.0 respectively.

It may be concluded that findings for these community youth groups are in agreement with hypothesis 6. It may be hypothesized additionally, through use of the r^2 as percentage explanation, that 64 to 75 per cent of the variance in sociometric scores is due to factors associated with the membership of different groups, such as the other five factors investigated here.

SUMMARY AND CONCLUSIONS

This investigation proposed to test the prediction for community youth groups of six hypotheses concerned with factors affecting sociometric acceptance and developed inductively from research conclusions. Ss were 270 girls in 23 Home Economics 4-H Clubs and 87 boys and 47 girls in eight Agriculture 4-H Clubs in the same county.

Four hypotheses predicted the relations found in these groups: (a) In child or youth groups, age has a positive, low correlation (.15-.20) with sociometric status. (b) Sociometric status in both sex groups is higher for girls than for boys. (c) Level of performance or achievement in the skills and knowledge essential for or being taught to the group is positively related to sociometric status. (d) Sociometric status in one group has a positive and fairly high correlation (.50-.60) with sociometric status in groups of different membership. One nonpredictive hypothesis specified a positive relation with leadership, defined here as election to club office. The other stated that family and community variables, defined here as farm or town residence and attendance at the same or different schools, may relate to social acceptance in particular groups.

Analyses of age differences indicated that in the girls' clubs deviations in age below the median age of the club were inversely related to sociometric scores, but deviations in age above the median age of the club did not relate to sociometric scores.

REFERENCES

1. BONNEY, M. E. Sociometric study of the relationship of some factors to mutual friendships on elementary, secondary and college levels. *Sociometry*, 1946, 9, 21-47.
2. BONNEY, M. E. Choosing between the sexes on a sociometric measurement. *J. soc. Psychol.*, 1954, 39, 99-114.
3. DAHLKE, H. O. Determinants of sociometric relations among children in the elementary school. *Sociometry*, 1953, 16, 327-338.
4. DAMRIN, DORA E. Family size and sibling age, sex and position as

- related to certain aspects of adjustment. *J. soc. Psychol.*, 1949, 29, 93-102.
5. GRONLUND, N. E. The relative stability of classroom social status with weighted and unweighted sociometric choices. *J. educ. Psychol.*, 1955, 46, 345-351.
 6. KOCH, HELEN L. A study of some factors conditioning the social distance between the sexes. *J. soc. Psychol.*, 1944, 20, 79-107.
 7. KOCH, HELEN L. The social distance between certain racial, nationality, and skin pigmentation groups in selected populations of American school children. *J. genet. Psychol.*, 1946, 68, 63-95.
 8. LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.
 9. LOOMIS, C. P., BAKER, W. B., & PROCTOR, C. The size of the family as related to social success of children. *Sociometry*, 1949, 12, 313-320.
 10. McCANDLESS, B. R., & MARSHALL, HELEN R. Sex differences in social acceptance and participation of preschool children. *Child Development*, 1957, 28, 421-425.
 11. MCGUIRE, C., MURPHY, A. C., & WHITE, G. D. Age-mate acceptance and adolescent assessments of one another. Research Memorandum No. 5, Laboratory of Human Behavior, Dept. of Educ. Psychol., Univer. of Texas, 1957.
 12. McNEMAR, Q. *Psychological statistics*. (2nd Ed.) New York: Wiley, 1955.
 13. MARSHALL, HELEN R. Some factors associated with social acceptance in women's groups. *J. Home Econ.*, 1957, 49, 173-176.
 14. MOUTON, JANE S., BLAKE, R. R., & FRUCHTER, B. The reliability of sociometric measures. *Sociometry*, 1955, 18, 7-48.
 15. MOUTON, JANE S., BLAKE, R. R., & FRUCHTER, B. The validity of sociometric responses. *Sociometry*, 1955, 18, 181-206.
 16. NEUGARTEN, BERNICE. Social class and friendship among school children. *Amer. J. Sociol.*, 1946, 51, 305-313.
 17. ORZACK, L. H. Preference and prejudice patterns among rural and urban schoolmates. *Rural Sociol.*, 1956, 21, 29-33.
 18. RADKE, MARIAN, SUTHERLAND, JEAN, & ROSENBERG, PEARL. Racial attitudes of children. *Sociometry*, 1950, 13, 154-171.
 19. SEEMAN, M. A. A situational approach to intra-group Negro attitudes. *Sociometry*, 1949, 9, 199-206.
 20. STENDLER, CELIA. *Children of Brasstown*. Urbana: Bureau of Research and Service, Coll. of Educ., Univer. of Illinois, 1949.
 21. THORPE, J. G. An investigation into some correlates of sociometric status within school classes. *Sociometry*, 1955, 18, 49-61.
 22. WITRYOL, S. L., & THOMPSON, G. G. A critical review of the stability of social acceptability scores obtained with the partial-rank-order and the paired comparison scales. *Genet. Psychol. Monogr.*, 1953, 48, 221-260.

DISCRIMINATION LEARNING IN CHILDREN AS A FUNCTION OF REINFORCEMENT VALUE

YVONNE BRACKBILL¹ AND DONALD JACK¹

San Jose State College

The present study is concerned with the influence of reinforcement *value* on learning. Reinforcement *value* is defined as the individual S's relative preference for various types of reinforcers, when amount of reinforcement and probability of occurrence of reinforcement are held constant.²

The results of several studies show that, on the average, different types of reinforcers are differentially effective for learning (1, 5). For example, Terrell and Kennedy found that children learn faster when rewarded with candy rather than with praise (5). However, despite the fact that children as a *group* learn more effectively on candy reinforcement, it is hardly conceivable that any one type and amount of reinforcement has an exactly equal reinforcing effect for all Ss of any such group. Given one group of Ss and one type of reinforcement, there must still be some intragroup variability in effectiveness of that reinforcement as a function of variability in value of the reinforcer for the Ss. In other words, for different people, the same objective or external reinforcer may have different reinforcement values, and if so, the within-group variability in reinforcement value should be reflected statistically as a large standard deviation in number of trials to learning criterion.

The present experiment used a discrimination learning situation to compare the relative effects of two experimental treatments: identity of reinforcement *value* vs. identity of external reinforcing object. Ss run under the first condition (group RV) were allowed to choose the most preferred one of three reinforcers: M & M candies, varicolored marbles, or varicolored plastic trinkets. Ss of the second experimental group (group R) were

¹ The authors wish to express their thanks and gratitude for the cooperation and assistance received from the following persons: Neal Royer, Superintendent, Campbell Elementary School District; Peter Franusich, School Psychologist; Duane Beaubien, Principal, Dover School; and Wayne Fontes, Principal, Hamilton School.

² From a conceptual point of view, *reinforcement value* might be considered as roughly analogous to *incentive motivation* (K), insofar as the latter intervening variable is used to account for performance differential as a function of differences in quality of reinforcement (3). There are operational differences, however, since the present procedure does not involve a consummatory response, the classical conditioning of which is assumed to be the basic mechanism underlying incentive motivation.

reinforced with candies. It was predicted that although there would be no significant mean difference between the two treatment groups in trials to criterion, group RV would show greater within-group homogeneity than group R, i.e., there would be a significantly smaller standard deviation for group RV.

Candy was used as the standard reinforcer for group R in order to make the experimental comparison a conservative one. The reasoning behind this was as follows: First, from a logical consideration, it would seem that if sigma were to vary in size at all with differences in *average* reinforcement value, it should be smallest for that reinforcer having the highest average value (and hence, the greatest facilitating effect on learning). [See, for instance, the relative sizes of standard deviations in Terrell and Kennedy's Table 1, (5, p. 259).] Second, from a statistical consideration, if means and standard deviations are positively correlated in size, that treatment producing lowest mean number of trials to criterion should also show the smallest standard deviation.

METHOD

Subjects

The Ss were 60 male children. They were obtained from the eight kindergarten classes of two elementary schools within a single school district. Mean age of Ss was 63.6 months, *SD*, 3.4. The sample was relatively homogeneous in terms of race (white) and socioeconomic status (middle class). Mean California Mental Maturity Scale IQs for second graders at each school was 106. Goodenough Draw-A-Man tests were available for all kindergartners, and were used in the selection process to eliminate potential Ss of subnormal intelligence, as follows: The school psychologist submitted a list of names of children with extremely low scores on the Goodenough test. A short-form WISC was then administered to these 17 children. Using a cutting score of 70 IQ, eight of the 17 were eliminated from the total pool from which Ss were drawn.

Selection of Ss and placement into one of the two experimental groups was in alphabetical order. In order to have the total RV group composed of equal numbers of Ss preferring each of the three reinforcers, a quota was established of 10 Ss per type of reinforcer. This quota imposed no practical difficulties, since, on the average, the three reinforcers were approximately equal in attractiveness. Only two potential Ss were discarded because they chose a reinforcer (trinkets) for which the quota had already been reached.

Experimental Procedure

S was seated opposite E, at a small table on which were three boxes, the discriminative stimuli. The boxes were of different colors (dark red, light red, and red-orange). Two boxes, No. 1 and No. 2, were the same

size but different in size from the third box.³ Color and size, however, were irrelevant cues, as was also the presence of box No. 3, since the correct response was a simple position alternation sequence of the reinforcement between boxes No. 1 and No. 2. For presentation, the boxes lay open side down on a large piece of brown felt approximately four inches distant from each other; the end boxes, No. 1 and No. 3, were equidistant from S. Position of the boxes was constant for all trials.

On a second table, to S's right, lay a small pillow on which S rested his head between trials. This served two purposes: (a) it eliminated the necessity of placing a screen in front of S while E reloaded the boxes, and (b) it reduced sound localization in the event of any auditory cues during reloading.

As soon as S entered the experimental room, E gave the following instructions:

(group R) Do you know what these are? (E paused and indicated an open box of M & M candies.) They're candy; they're chocolate. I have something very special for you to do, and if you can do it you can have some of the candy.

(group RV) Here are some marbles; here are some charms; here are some candies. (E paused and indicated the appropriate boxes; placement and naming order were rotated by Ss.) Now I have something very special for you to do, and if you can do it you can have some of the charms OR some of the candies OR some of the marbles. You can't have some of all of them, just some of one kind. Now THINK HARD: if you could have JUST ONE KIND, which kind would you rather have—would you rather have some candies OR would you rather have some charms OR would you rather have some marbles?

(Both groups) I'm going to put one — at a time under one of these boxes. And every time you pick up the box that has the — under it, you can keep the —. Understand? Now! THERE IS A WAY TO FIND THE RIGHT BOX EVERY SINGLE TIME. SEE IF YOU CAN FIND THE WAY.

Put your head down on that pillow and wait until I say "ready." . . . Ready. Pick up the box that you think has the — under it. . . . Now every time you get a —, pick it up and put it in this envelope, and all the — that you put in there will be for you to keep. . . . Head down. . . . Ready.

(After the last trial, the following statement was repeated until S indicated agreement. The purpose of the statement was to "jam" inter-S communication.)

Well, you found the way, didn't you? You found out that the candy was always under the RED box. (E indicated simultaneously all boxes.) Yes, the candy was always under the RED box.

³ The boxes were of ordinary heavy cardboard construction. Box No. 1 measured $4\frac{1}{4}$ by $3\frac{1}{2}$ by $1\frac{1}{8}$ in., and was painted with a mixture of 2 parts red Dope to 1 part white Dope. Box No. 2 had the same dimensions as Box No. 1, and was painted with a mixture of 1 part red Dope to 2 parts white Dope. Box No. 3 measured $3\frac{1}{8}$ by $2\frac{1}{4}$ by $1\frac{3}{8}$ in., and was painted with orange Dope.

CHILD DEVELOPMENT

Two learning criteria were used: (a) number of trials taken before *S* no longer chose the irrelevant or never-reinforced box No. 3; and (b) the number of trials taken to a run of 10 consecutive correct choices. An experimental session ended either when *S* reached the second criterion or after the 150th trial. In the latter case, the score assignment was 150. Session length ranged from 10 to 25 minutes per *S*, depending upon time taken to reach the second criterion.

TABLE I
MEANS AND VARIANCE ESTIMATES FOR DISCRIMINATION
LEARNING CRITERIA

	<i>First Criterion: number of trials to eliminate irrelevant box</i>		<i>Second Criterion: number of trials to and including 10 consecutive correct responses</i>	
	G R O U P		G R O U P	
	R	RV	R	RV
<i>N</i>	30	30	30	30
<i>M</i>	58.30	43.57	72.47	60.97
σ_{DM}		11.15		10.55
<i>t</i>		1.32		1.09
<i>p</i>		.19		.27
σ^2	2503.96	1351.25	2283.13	1167.34
<i>F</i>		1.85		1.96
<i>p</i>		.05*		.05*

* One-tailed test.

RESULTS

Table I shows the means and standard deviations for both groups on both learning criteria. As predicted, there were negligible differences between means and relatively large differences between standard deviations. That is, on both measures, the RV group showed significantly less variability than did the R group. Although the evidence is by no means overwhelming in support of the hypothesis, it ought to be pointed out that the experimental comparison was a conservative one in two ways. The first—the use of a high average value reinforcer for group R—was mentioned above. Second, the true variability was underestimated by assigning to nonlearners the limiting score of 150. This restriction worked against the hypothesis since there was a disproportionately larger number of nonlearners in group R than in group RV (six as compared to one).

The correlation between CA and number of trials to the second criterion was $-.12$. The insignificant size of this correlation is no doubt

attributable to restriction of age range in the sample. Last, there was no significant difference in mean performance between the 10 candy reinforced Ss of group RV and the 30 candy reinforced Ss of group R.

DISCUSSION

The results of the present study are relevant to some common methodological problems. First, an experimenter who employs a simple analysis of variance design, using children as Ss frequently finds that between-group differences are large, but that the size of the within-group variability is even more impressive. The present findings suggest that the use of individually determined reinforcers will reduce such error variance.

Second, experimenters working with young children often report some difficulty in maintaining sufficient motivation to ensure Ss' cooperation or even to keep them in the experimental room. The method of allowing S to choose that reinforcer he would most like to work for probably maximizes motivation as a function of high incentive value.

A third methodological problem concerns the fact that the effectiveness of any given type of reinforcer varies with age. For example, Sturgis (4) has reported that trinkets are not effective reinforcers at age $2\frac{1}{2}$, but are effective at age 4. What is needed, then, is a "methodological cookbook" of normative data, listing relative effectiveness of all possible reinforcers by successive age levels. But this information is not available, except for the limited data provided by Terrell and Kennedy. At present, selection of the most effective reinforcer for a given problem and age group depends on a good guess or a pilot study. The present results suggest that it would be less hazardous and/or time-consuming to hold reinforcement value constant for any experimental group rather than reinforcement type, i.e., to offer Ss a selection of reinforcers.

The above data may also lend themselves to practical application. For example, in the classroom situation, teachers depend principally on one type of reinforcement—the grade. However, grades are not equal in reinforcement value for all students. In view of this, teachers might obtain better academic results from their poorer students by offering other, individually determined, incentives. As an incidental note, one of the authors has tried informal remedial training with a few children deficient in either arithmetic or reading. In all cases, the under-achievement appeared to be a function of insufficient motivation, not of inability. Rapid improvement was noted when the procedure included (a) reinforcement with individually determined reinforcers, and (b) a combined reward-punishment technique (2).

SUMMARY

A comparison was made of the relative effects on discrimination learning of two experimental treatments: identity of reinforcement value vs.

CHILD DEVELOPMENT

identity of external reinforcing object. Reinforcement value was defined as S's relative preference for various types of reinforcers, with amount of reinforcement and probability of occurrence of reinforcement held constant. Results supported the expectation that, although the two groups would not differ in mean trials to criterion, the reinforcement value group would show significantly less variability in trials to criterion.

REFERENCES

1. ABEL, L. B. The effects of shift in motivation upon the learning of a sensori-motor task. *Arch. Psychol.*, N.Y., 1936, 29, No. 205.
2. BRACKBILL, YVONNE, & O'HARA, JOHN. Discrimination learning in children as a function of reward and punishment. Paper read at Western Psychol. Ass., Eugene, Oregon, 1957.
3. SPENCE, K. W. *Behavior theory and conditioning*. New Haven: Yale Univer. Press, 1956.
4. STURGIS, PERSIS T. What are some of the positive reinforcers suitable for preschool children? Paper read at Western Psychol. Ass., Eugene, Oregon, 1957.
5. TERRELL, GLENN, JR., & KENNEDY, W. A. Discrimination learning and transposition in children as a function of the nature of the reward. *J. exp. Psychol.*, 1957, 53, 257-260.

NURTURANCE AND NURTURANCE-WITHDRAWAL IN RELATION TO THE DEPENDENCY BEHAVIOR OF PRESCHOOL CHILDREN¹

WILLARD W. HARTUP

Iowa Child Welfare Research Station

This investigation is based on the hypothesis that non-nurturance by an adult is more strongly associated with the occurrence of dependency behavior in young children than is nurturance alone. The naturalistic studies of Sears *et al.* (11), Beller (1), and Smith (12) all contain data which show a positive relationship between amount of parental frustrations (non-nurturance) and the frequency of dependency behavior observed in young children. Similar results were obtained in the laboratory studies of Gewirtz (5, 6, 7), although Carl's laboratory findings with respect to this hypothesis were inconclusive (2).

The present study was designed to explore the relationship between one specific form of non-nurturance—the withdrawal of nurturance—and young children's acquisition of responses which elicit adult approval. This relationship was studied in the laboratory where some manipulation of the relevant antecedent conditions was possible.

The method of this study has been to provide a comparison in the learning of simple responses which elicit adult approval between a group of children consistently nurtured by an experimenter and another group who were nurtured and then rebuffed (nurturance-withdrawal). It was predicted that children in the presence of an adult female experimenter who withdraws her nurturance in this fashion will learn simple tasks eliciting adult approval in fewer trials and with fewer errors than children

¹ This study is based on a doctoral dissertation submitted by the author to the Graduate School of Education of Harvard University. Parts of this investigation were reported at the 1956 meetings of the American Psychological Association in Chicago, Illinois. The author wishes to acknowledge the helpful criticisms of Dr. Harry Levin in the preparation of this study.

in the presence of an experimenter who has been consistently nurturant. This prediction is based, in part, on those aspects of psychoanalytic theory which suggest that attempts by the child to institute closeness and seek affection are most strongly related to the anxiety generated at times of separation from the mother or when the child has experienced loss of the mother's love (4). It is believed that the withdrawal of nurturance by a female experimenter is similar to certain aspects of the caretaker-child relationship. If so, such behavior by the experimenter should generate certain amounts of anxiety in young children which, in turn, should motivate dependency behavior.

Certain aspects of behavior theory are also relevant to the present prediction. Miller (9) and Mowrer (10) suggest that the capacity of a neutral stimulus to evoke anxiety is strengthened through association with increases in drive or delay in primary reinforcement; presumably it is association such as this which results in children becoming anxious when the mother is absent or non-nurturant. If children in our culture commonly do learn to respond in this way to the non-nurturance of adults, if anxiety does motivate behavior, and if adult nurturance has acquired the capacity to reduce anxiety and thereby reinforce behavior for the young child, the experimental prediction formulated for this investigation can be made.

METHOD

Subjects

Subjects used in this investigation were 34 preschool children—15 boys and 19 girls—in attendance at the Harvard University Preschool during the spring of 1954. The subjects ranged in age from three years, ten months, to five years, six months. The mean age of the children in the sample was four years, seven months; the standard deviation was 4.7 months. The preschool population at the time of the experiment was typical of many laboratory nursery schools in that the children were all from academic, professional, or business homes, and were free from severe emotional disturbances and physical handicaps. Two subgroups were drawn from this sample. These subgroups were counterbalanced with respect to sex of child and dependency ratings made by the preschool teachers on scales of the type used by Beller (1). Two young women served as experimenters.² Each experimenter worked with a randomly-assigned half of the subjects in each group.

Procedure

Each child was brought individually to the laboratory room for the experimental session. This room was equipped with one-way mirrors for observation and was furnished with a child's table and chairs, an adult-

² The author wishes to thank Mrs. Carla F. Berry and Miss Willa Dinwoodie for their assistance in this regard.

sized table with comfortable chair, and a large bench. The experimental session proceeded as follows:

1. For a period of five minutes, the experimenter interacted nurturantly with the child while the child played with toys. For purposes of this experiment nurturance consisted of adult behavior which rewarded, encouraged, supported, or showed affection to the child; during this five-minute period the experimenter attempted to maximize these qualities in her behavior toward the child. Children in both experimental groups experienced this period of nurturant interaction with the experimenter.

2. Children in the consistent-nurturance group (hereafter called group C) then immediately experienced a second five-minute period like the first.

3. The second five minutes for the nurturance-withdrawal group (group NW) were marked by the experimenter's behaving non-nurturantly toward the child. She ceased to interact with the child, withdrew from his proximity, and did not reward any of the child's supplications beyond telling him that she was "busy." The experience of children in group NW, having first a period of nurturant interaction, then a period of non-nurturance from the experimenter, has been called "nurturance-withdrawal."

4. Children in both experimental groups were then asked by the experimenter to learn two tasks, the reward for which was the verbal approval of the experimenter. Task I consisted of learning a simple *position* concept in an arrangement of two blue and two red one-inch blocks. The task was presented to the child as a guessing game. The experimenter placed the blocks on the floor first in this order (reading from the child's left): red, red, blue, blue. She then said: "I'm thinking of one of the blocks and I want to see if you can guess which one it is. Point with your finger to the one you think is right and I'll tell you if it's the right one." The child's first guess was always unsuccessful, as was his second. The third guess was always successful. This introductory procedure was followed to eliminate chance successes on the first guess. On each succeeding trial the arrangement of the blocks was changed through all the possible order-permutations. The correct block was always the block in the same position in the row as the one which the child chose on his third guess. The performance criterion was three consecutive correct trials. Task II consisted of copying from memory a row of adjacent blue, red, and yellow one-inch cubes which were shown to the child for five seconds per trial. Six blocks were arranged in the following order: red, yellow, blue, blue, yellow, red. The performance criterion was one perfect reproduction of the arrangement completed by the child from his own supply of blocks. Measures used in the subsequent analysis of the data were: (a) number of errors to criterion on task I; (b) number of trials to criterion on task I; (c) number of errors to criterion on task II; (d) number of trials to criterion on task II. Error- and trial-scores were correlated .93 on task I, .96 on task II.

Measures of the child's tendency to be dependent on adults were from three sources: (a) observation during the period of nurturant interaction

CHILD DEVELOPMENT

in the laboratory; (b) ratings of the child's dependency on adults made by the preschool teachers; and (c) observations of the child's dependency on adults in preschool made by observers.

Behavior categories used for the laboratory observations were as follows:

1. Asks for verbal help and information
2. Asks for material help
3. Seeks reassurance and rewards
4. Seeks positive attention
5. Seeks to be near
6. Seeks physical contact
7. Seeks negative attention
8. Initiates verbal interaction with experimenter

Frequencies in categories 1 through 7 were summed to yield a measure of dependence on the experimenter. Category 8, "verbal interaction," was used independently in the analysis of data.

The following seven-point scales were used for the teacher ratings of the child's dependence on adults in the preschool situation:

1. Seeks recognition
2. Seeks unnecessary help
3. Seeks necessary help
4. Seeks physical contact and proximity
5. Seeks attention

Each child was rated by two of his teachers (reliability coefficients ranged between .73 and .99). The ratings of the two teachers on each scale were pooled; a summary rating score was then obtained by summing the pooled ratings on all five scales. This summary rating score was used to counterbalance the two subgroups as described above and was also used in analyzing the learning scores.

Behavior categories used for the preschool observations of dependency on the teacher were:

1. Seeks recognition and approval
2. Seeks unnecessary help
3. Seeks necessary help
4. Seeks physical contact
5. Seeks to be near
6. Seeks positive attention
7. Seeks negative attention

Frequencies in the seven categories were summed to yield the preschool observation measure of dependence on the teacher which was used in the statistical analysis.

The intercorrelations among the two laboratory scores, the summary teacher rating, and the preschool observation total score are reported in Table 1.

TABLE 1

INTERCORRELATIONS AMONG FOUR MEASURES OF DEPENDENCY ($N = 34$)

	1	2	3	4
1. Verbal interaction with experimenter: laboratory	—			
2. Dependence on experimenter: laboratory	.13	—		
3. Dependence on preschool teacher: teacher rating	.03	.40*	—	
4. Dependence on preschool teacher: observer score	.00	.11	.31	—

* Significant beyond .05 level.

The data from the learning tasks were studied by a triple-classification analysis of variance technique for unequal cell-entries (13). The three independent variables in this analysis were (a) sex of child; (b) dependency scores as described above and by which the group was separated into two subgroups—high dependency (all cases above the median on the score being used) and low dependency (all cases below the median); (c) experimental treatment, consistent nurturance versus nurturance-withdrawal. Four analyses of variance were completed for each trial- or error-score from the learning tasks: one for each summary dependency score described above. The results of these analyses are summarized in the tables which follow.

TABLE 2

F RATIOS FROM FOUR ANALYSES OF VARIANCE BASED ON NUMBER OF TRIALS ON TASK 1 ACCORDING TO SEX OF CHILD, DEPENDENCE, AND EXPERIMENTAL CONDITION

Source	Analysis 1* ($N=27$)	Analysis 2* ($N=33$)	Analysis 3* ($N=32$)	Analysis 4* ($N=31$)
Sex of Child	16.153 [§]	2.618	3.988	2.953
Dependence	7.163 [†]	.002	.527	6.571 [†]
Experimental Condition	13.859 [‡]	4.574 [†]	1.759	1.438
Sex × Dependence	3.053	3.630	.028	.015
Sex × Condition	6.744 [†]	2.098	3.549	6.437
Dependence × Condition	2.224	.792	2.249	1.488
Sex × Dependence × Condition	.215	.570	.005	1.186

* The measures of dependence used were: analysis 1, frequency of verbal interaction initiated by child in the laboratory session; analysis 2, frequency of dependence on adults observed in the laboratory session; analysis 3, teachers' ratings of dependence; analysis 4, dependence on preschool teachers as recorded by observers.

† Significant between .05 and .01.

‡ Significant between .01 and .001.

§ Significant beyond .001.

CHILD DEVELOPMENT

RESULTS

Task I—Number of Trials

Table 2 shows that the F ratio for the experimental variable was significant beyond the .01 level in analysis 1 and beyond the .05 level in analysis 2. The means for groups C and NW are reported in Table 3 along with the t ratio for the difference between means. These data show that NW children (who had experienced nurturance-withdrawal) took fewer trials to complete the task than C children (who had experienced consistent nurturance). Table 2 also shows that in two analyses there was a significant interaction between experimental condition and the sex of child. This interaction may be interpreted to indicate that the effects of nurturance-withdrawal by a female adult are dependent on the sex of the child. The data reported in Table 3 show more clearly the effects of this interaction: nurturance-withdrawal was clearly associated with faster learning of the task for girls, but there was no difference between the means for boys in group C and in group NW.

TABLE 3
MEAN NUMBER OF TRIALS ON TASK I FOR BOYS AND GIRLS
IN TWO EXPERIMENTAL GROUPS

	Group C	N	Group NW	N	t	p
Boys	15.86	7	16.14	7	..	n.s.
Girls	35.22	9	16.10	10	2.570	.02 > p > .01
Boys and Girls	26.77	16	16.11	17	2.041	.05 > p > .01

The F ratio for sex of child was significant in one of the analyses reported in Table 2, and the direction of mean differences showed that boys as a group learned the task in fewer trials than girls. This difference is believed to relate to some feature of the task itself rather than the social conditions of the experiment since the same difference was not found in the data for task II; however, the aspect of task I producing the sex difference was not clear from observation of the children in the experiment. The dependency variable was also significant in two analyses; subgroup means showed that the more dependent children (as measured by verbal interaction) learned task I more quickly than the less dependent children, but that more dependent children (as measured by preschool observation) learned task I less quickly than the less dependent children.

Task I—Number of Errors

Since the error-measure was highly correlated with the trials-measure on task I similar results would be expected and were obtained from the analyses of variance. F ratios for the experimental variable were significant in analyses 1 and 2 (Table 4). Again, significant interaction between sex

TABLE 4

F RATIOS FROM FOUR ANALYSES OF VARIANCE BASED ON NUMBER OF ERRORS ON TASK I ACCORDING TO SEX OF CHILD, DEPENDENCE, AND EXPERIMENTAL CONDITION

Source	Analysis 1* (N=27)	Analysis 2* (N=33)	Analysis 3* (N=32)	Analysis 4* (N=31)
Sex of Child	9.421†	1.360	2.581	1.385
Dependence	5.060†	.190	1.050	6.577†
Experimental Condition	10.519†	4.640†	1.232	.924
Sex × Dependence	1.943	4.063	.003	.532
Sex × Condition	4.267	5.560	3.262	6.118†
Dependence × Condition	2.458	1.364	1.610	.772
Sex × Dependence × Condition	.074	1.219	.003	1.414

* The measures of dependence used were: analysis 1, frequency of verbal interaction initiated by child in the laboratory session; analysis 2, frequency of dependence on adults observed in the laboratory session; analysis 3, teachers' ratings of dependence; analysis 4, dependence on preschool teachers as recorded by observers.

† Significant between .05 and .01.

‡ Significant between .01 and .001.

of child and experimental condition was found (in analysis 4, Table 4). Table 5 reports the means for the experimental groups according to sex of child, and once more the results show significant differences between the experimental groups for girls but not for boys.

The *F* ratio for sex of child was significant in one analysis indicating that boys learned the task faster than girls. The dependency variable was significant in two analyses, but mean differences were in the same inconsistent directions as in the analysis of trial-scores on task I.

TABLE 5
MEAN NUMBER OF ERRORS ON TASK I FOR BOYS AND GIRLS
IN TWO EXPERIMENTAL GROUPS

	Group C	N	Group NW	N	t	p
						n.s.
Boys	18.14	7	18.71	7	2.71	.02 > p > .01
Girls	39.33	9	17.30	10	1.90	.10 > p > .05
Boys and Girls	30.06	16	17.88	17		

Task II—Number of Trials

The distribution of scores on this measure significantly departed from normality; hence the results are not reported here.

Task II—Number of Errors

The results for errors on task II were consistent with respect to experimental condition with those found on task I, although at a lesser level of

TABLE 6

F RATIOS FROM FOUR ANALYSES OF VARIANCE BASED ON NUMBER OF ERRORS ON TASK II ACCORDING TO SEX OF CHILD, DEPENDENCE, AND EXPERIMENTAL CONDITION

Source	Analysis 1* (N=24)	Analysis 2* (N=29)	Analysis 3* (N=28)	Analysis 4* (N=28)
Sex of Child	.169	1.567	.170	.056
Dependence	2.146	.493	5.318†	1.359
Experimental Condition	.501	2.788	4.755†	.521
Sex × Dependence	2.332	1.691	2.667	.314
Sex × Condition	.003	.014	.373	1.682
Dependence × Condition	1.155	.037	.311	.167
Sex × Dependence × Condition ..	.263	.048	.376	1.983

* The measures of dependence used were: analysis 1, frequency of verbal interaction initiated by child in the laboratory session; analysis 2, frequency of dependence on adults observed in the laboratory session; analysis 3, teachers' ratings of dependence; analysis 4, dependence on preschool teachers as recorded by observers.

† Significant between .05 and .01.

significance. Table 6 shows that one *F* ratio for the conditions variable was significant beyond the .05 level.

The mean number of errors on task II made by the two experimental groups are reported in Table 7. These data are consistent with those for task I and suggest that faster learning was produced under the nurturance-withdrawal condition than under a condition of uninterrupted nurturance. No significant interaction effects were discovered in the analyses of variance.

The *F* ratios for sex of child were not significant in these analyses; the dependency variable (when measured by teacher ratings) yielded a significant *F*, and the means suggest that more dependent children learned task II faster than less dependent children.

TABLE 7
MEAN NUMBER OF ERRORS ON TASK II FOR BOYS AND GIRLS
IN TWO EXPERIMENTAL GROUPS

	Group C	N	Group NW	N	<i>t</i>	<i>p</i>
Boys	12.40	5	6.17	6	1.121	.30 > <i>p</i> > .20
Girls	8.33	9	4.44	9	1.420	.20 > <i>p</i> > .10
Boys and Girls	9.79	14	5.07	15	1.750	.10 > <i>p</i> > .05

DISCUSSION

The findings for girls uniformly support the hypothesis that nurturance-withdrawal is associated with more efficient performance on the learning tasks than consistent nurturance. The results for boys, however, showed that there were no differences between the nurturance-withdrawal and the consistent nurturance groups. Actually, the results for boys were not so clearly negative. When the boys' groups were divided according to the measures of dependence, *highly* dependent boys were found to respond much as the girls while *low* dependent boys responded in the reverse fashion. Thus, highly dependent boys (who may be assumed to be generally anxious concerning their relationships with adults) did learn more efficiently when the experimenter withdrew her nurturance. The boys in the low dependency group who were consistently nurtured learned more efficiently than boys in this group who experienced nurturance-withdrawal. Although the number of cases in these subgroups was small, this trend in the data suggests support for the hypothesis concerning the influence of nurturance-withdrawal for highly dependent boys as well as for the girls.

The fact that the findings for boys were so equivocal is an interesting one for further exploration. It may be that boys respond differently from girls to the nurturance-withdrawal of a *female* experimenter. Psychoanalytic theory regarding Oedipal relationships suggests that the punitiveness of a like-sexed adult should be more threatening to a child than the punitiveness of an opposite-sexed adult. It would be comparatively easy to incorporate sex of experimenter as a variable in a study such as this. However, since the results of this study suggest that nurturance-withdrawal fails to motivate only *low* dependent boys, there may be some sort of complex interaction among experimental conditions which this study failed to bring to light. For example, boys of this age who are not overtly dependent on adults may have moved further than highly dependent boys toward identifying with the male sex role, which in our culture contains certain elements of independence and self-reliance. Boys who *are* highly identified with the male role might well respond with greater anxiety to the mothering nurturance of the experimenter than to the condition in which the experimenter ceases to be attentive and leaves the child alone. If this were the case, most efficient learning would then have taken place in the low dependency boys under conditions of consistent nurturance. The experiment contains no measure of identification by which this interpretation can be checked.

The relationships of sex of child and dependence to speed of learning on the experimental tasks are far from being clear-cut in these data. Boys tended to learn task I faster than girls, but not task II. Dependence, as measured either by verbal interaction in the laboratory session or by teacher ratings, was positively associated with speed of learning; however, dependence as rated by observers in the preschool proved to be negatively asso-

ciated with speed of learning. The sources of these inconsistencies probably lie in the situational and procedural differences involved in the various measures of dependence; the intercorrelations among the dependency measures suggest that the various measures did *not* measure similar behavioral traits. It does not appear possible, therefore, without further study, to suggest the manner in which motivation to be dependent modifies the effects of nurturance-withdrawal on speed of learning in young children.

The general significance of these findings is felt to be clear and in the direction suggested by the hypothesis tested: nurturance-withdrawal stimulates faster learning than nurturance alone on simple cognitive tasks for girls, and probably also for boys. There may be, however, some second- or third-order interaction between nurturance-withdrawal, sex of child, sex of experimenter, and dependence which influences the behavior of boys under conditions like those of this experiment.

SUMMARY

Thirty-four four-year-old preschool children were divided into two experimental groups equated on the basis of sex and teachers' ratings of dependence on adults in preschool. Individual subjects in one group (C) were consistently nurtured by a female experimenter during a 10-minute period of interaction, after which two simple tasks were learned by the child. The subjects in the second group (NW) experienced nurturant interaction with the experimenter during only five minutes, then experienced five minutes of non-nurturant response from the experimenter, and finally were asked to learn the tasks. The data were treated by an analysis of variance technique in which learning scores were divided according to sex of child, dependency ratings, and experimental condition.

Children in group NW took fewer trials to learn task I than children in group C. Although the group findings were significant, analysis according to sex of child showed that nurturance-withdrawal was most clearly associated with faster learning in girls. Children in group NW made fewer errors in learning both task I and task II, although these findings were most significant for girls.

It is felt that these results support the hypothesis that nurturance-withdrawal supplies greater motivation than consistent nurturance for children's behavior which is designed to gain the reassurance of adults.

REFERENCES

1. BELLER, E. K. Dependency and independence in young children. *J. genet. Psychol.*, 1955, 87, 25-35.
2. CARL, J. An experimental study of the effect of nurturance on preschool children. Unpublished doctoral dissertation, State Univer. of Iowa, 1949.

WILLARD W. HARTUP

3. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
4. FREUD, S. Three contributions to the theory of sex. In *The Basic Writings of Sigmund Freud*. New York: Random House, 1938.
5. GEWIRTZ, J. L. Succorance in young children. Unpublished doctoral dissertation, State Univer. of Iowa, 1948.
6. GEWIRTZ, J. L. Three determinants of attention-seeking in young children. *Monogr. Soc. Res. Child Developm.*, 1954, 19, No. 2. (Serial No. 59)
7. GEWIRTZ, J. L. Does brief social deprivation enhance the effectiveness of a social reinforcer ("approval")? *Amer. Psychologist*, 1956, 11, 428. (Abstract)
8. HARTUP, W. W. Nurturance and nurturance-withdrawal in relation to the dependency behavior of preschool children. Unpublished doctoral dissertation, Harvard Univer., 1955.
9. MILLER, N. E. Studies of fear as an acquirable drive. I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. exp. Psychol.*, 1948, 38, 89-101.
10. MOWRER, O. H. *Learning theory and personality dynamics*. New York: Ronald, 1950.
11. SEARS, R. R., *et al.* Some child-rearing antecedents of dependency and aggression in young children. *Genet. Psychol. Monogr.*, 1953, 47, 135-234.
12. SMITH, H. T. A comparison of interview and observation measures of mother behavior. Unpublished doctoral dissertation, Harvard Univer., 1953.
13. WALKER, H. M., & LEV, J. *Statistical inference*. New York: Holt, 1953.

TEACHERS' AND PEERS' PERCEPTIONS OF CHILDREN'S SELF-CONCEPTS¹

HUGH V. PERKINS

*Institute for Child Study
University of Maryland*

In recent years efforts which have been directed toward achieving a more adequate interpretation of behavior have centered in the development and further refinement of self theory. The growing importance of the self-concept as a construct is particularly evident in those theories which postulate that the individual's perception of himself is the central factor influencing his behavior (6, 7, 11, 13). These formulations suggest that a more adequate interpretation of behavior can be achieved when the observer increases his knowledge of the behavior's perceptual field including his self-concept.

The utilization and further development of this principle has been especially widespread in the areas of counseling and psychotherapy. In numerous researches the self-concept has proved to be an especially promising measure for evaluating changes which occur during therapy (2, 3, 12). The interest of educators, too, has increasingly focused on the self-concept, for it appears to be a valuable tool in the further investigation of psychological factors which influence learning and development (4, 8). Specifically, the self-concept can be used in education (a) as a psychological construct which enables teachers, counselors, parents and others to achieve with training deeper understandings and insights into the behavior and development of children, and (b) as a vital and important aspect of learn-

¹ The writer desires to acknowledge the services and assistance of many persons who contributed to this research study. He is especially appreciative of the help and support given him by Drs. Howard A. Lane, Louise Antz, Mary Beauchamp, and Louis Rath, all of New York University. He is indebted to Mr. William S. Schmidt, Miss Rowenetta Allen, and Mrs. Catherine Reed for their help in the initial phases of the study. To the principals, teachers, and children whose participation and cooperation made this study possible, the writer is especially grateful. Finally, appreciation is extended to the writer's colleagues, the staff members of the Institute for Child Study of the University of Maryland, for their encouragement and many helpful suggestions, and to the Grant Foundation for their generous financial support.

ing and development which the school through its educational processes seeks to promote and foster in every child.

An investigation was made of selected factors influencing perceptions of and changes in children's self-concepts in order that the uses of the self-concept in education might be fully ascertained. In addition, the study sought to explore the further implications of this construct for educational practice. This article will report that part of the study which investigated the effects of selected factors on perceptions of children's self-concepts; the second will concern the effects of selected factors on changes in children's self-concepts.

THE PROBLEM

Specifically, then, the first part of this study investigated the effects on teachers' and peer groups' perceptions of children's self-concepts in the fourth and sixth grades of four factors: (a) social-emotional climate, (b) teacher participation in an in-service child study program, (c) teacher acceptance of self and others, and (d) grade level.

Self-concept in this study was defined as those perceptions, beliefs, feelings, attitudes, and values which the individual views as part or characteristic of himself. In this study evidences of teachers' perceptions of and children's expressed self-concepts were obtained by means of a Q-sort, an instrument and technique which requires the subject to make a forced normal distribution of a series of self-referent statements.

Social-emotional climate was defined as the feeling tone or quality of interpersonal feeling created by group interaction. Classroom climate was defined operationally in this study as the ratio of teacher statements categorized as learner-supportive to those categorized as teacher-supportive as determined by Withall's Index (15).

Participation in child study was defined as teacher participation for three years in the teacher in-service program in child study sponsored by the Institute for Child Study of the University of Maryland. In the third year of this in-service study program teachers seek a deeper understanding of behavior through making hypotheses of children's self-concepts in the many situations described in their case records on these children.

Teacher acceptance of self and others was defined as the attitudes of valuing and concern which a teacher reveals in relation to herself and others in her environment. In this study the results of Reed's Sentence Completion Test (10) were used to identify teacher differences in acceptance of self and others.

HYPOTHESES OF THE STUDY

The general hypothesis which was tested by the first part of this study is as follows: In their perceptions of children's self-concepts teachers in group-centered climates, those who have completed three years of child study, those who are more accepting of self and others, and those who teach fourth grade will show greater similarity to these children's expressed self-concepts than will, respectively, teachers in teacher-centered climates,

those who have never participated in child study, those who are less accepting of self and others, and those who teach sixth grade. The general hypothesis for peer groups' perceptions of children's self-concepts is similar except for the variable of grade level where peer groups of sixth grade are favored over those of the fourth grade.

In addition to the general hypotheses stated above the following specific hypotheses were also tested:

1. Teachers' perceptions of children's self-concepts are in general positively and significantly related to these children's expressed self-concepts.
2. Teachers change significantly toward achieving greater correspondence between their perceptions of children's self-concepts and these children's expressed self-concepts.
3. Sixth grade peer groups' perceptions of sixth grade children's self-concepts will be more similar and fourth grade peer groups' perceptions of fourth grade children's self-concepts will be less similar than the expressed self-concepts of the respective sixth and fourth grade children.
4. Peer groups as a whole will show for "less threatening" percepts of children's self-concepts a greater similarity to those children's expressed self-concepts than they will show for "more threatening" percepts of children's self-concepts.

THE SAMPLE

A representative sample of eight classrooms in seven elementary schools which are part of a large county school system in Maryland was selected for study. The communities served by these schools are suburban and are predominantly lower middle to upper middle class in socioeconomic level. Classrooms were selected on the basis of their being willing to participate in the research and their being representative of one of the extremes with respect to each of the independent variables: (a) social-emotional climate, group-centered *vs.* teacher-centered; (b) participation in child study, completion of the three-year program in child study *vs.* no participation in program; (c) teacher acceptance of self and others, more acceptance *vs.* less acceptance of self and others; and (d) grade level, fourth *vs.* sixth. The factorial design of this research in which extremes of each of the four independent variables are represented in the eight classrooms of the sample is shown in Figure 1.

So few fourth and sixth grade teachers who had completed three years of child study were available that nearly all of the classrooms taught by these teachers were included in the study. Assessment of the social-emotional climate of these classrooms and others of the sample was made by the investigator and another observer using Withall's Climate Index (15). This instrument contains seven descriptive categories which form a learner-centered to teacher-centered continuum, and the procedure consists of cate-

CHILD DEVELOPMENT

		GRADE LEVEL				
		Fourth		Sixth		
CLIMATE	Group-Centered	A More	J More	D Less	L Less	TEACHER ACCEPTANCE OF SELF AND OTHERS More Less
	Teacher-Centered	B Less	K Less	C More	M More	
		CS	NCS	CS	NCS	
CHILD STUDY PARTICIPATION						

CHILD STUDY PARTICIPATION

FIGURE 1—Factorial design of research showing distribution of differences on four independent variables among the eight classrooms of the sample.

gorizing teacher statements made in classroom interaction. A measure of classroom climate is obtained by contrasting the proportion of statements categorized as learner-supportive (group-centered) with the proportion categorized as teacher-supportive (teacher-centered). The objectivity of the investigator's use of the Climate Index was established by his obtaining a mean agreement of 88 per cent with a second observer for independent categorizations of the same statements of teachers in two classrooms. A reassessment of classroom climate using the Climate Index near the end of the study showed that the climate in these classrooms had remained stable except for one classroom which had had a change in teachers near the end of the study.

Teachers of classrooms in the sample were designed as "more" or "less" acceptant of self and others in accordance with the scores each made on Reed's Sentence Completion Test (10) which measures teacher acceptance of self and others. The four teachers making the highest scores on Reed's test were judged to be more acceptant and the four teachers with low scores were judged to be less acceptant of self and others. In scoring the test the mean percentage of agreement between the investigator and a colleague who had had considerable experience in using and scoring the test was 82 per cent.

The eight teachers of the sample varied in teaching experience from 3 to 33 years and in education from those having no college degree to one with a master's degree. A previous study suggests that these differences in teaching experience and professional training are not factors which would influence significantly the results of this study. Kindred (5) found that in their use of human development principles teachers with additional graduate hours were not significantly different from teachers with less education. This was true also for teachers who differed with respect to number of years' teaching experience.

DEVELOPMENT OF INSTRUMENTS

The principal instrument used in this research to study children's self-concepts and teachers' perceptions of children's self-concepts was the Q-sort. The Q-sort, a part of the Q-technique developed by Stephenson (14), is a method of gathering and scoring items for purposes of correlating people instead of tests.

The Q-sort developed for use in this study was a stratified random sample of 50 self-referrent statements, selected from a universe of statements which had been derived from responses of fourth and sixth grade children reported in a survey by Jersild (4). This investigator used the items and subpoints listed under each of Jersild's categories to develop self-referrent statements using language which he believed would be generally characteristic of fourth and sixth grade children. The following 50-item Q-sort was constructed by selecting at random from each category the number of self-referrent statements that corresponded with the frequency of self-percepts that were mentioned in that category by fourth and sixth grade children in Jersild's study.

1. I look on the bright side of things.
2. I understand the kind of a person I am.
3. I am a fast runner.
4. I am not neat.
5. I can take a joke on myself.
6. I cannot talk well in front of a group.
7. I join in doing what the group is doing.
8. I am a leader.
9. My clothes are different than the kind other people wear.
10. I do not like to make things.
11. I keep working until my work is finished.
12. I don't like arithmetic.
13. I don't like animals.
14. I get excited and upset easily.
15. Other people want me to tell them what to do.
16. I draw pictures.
17. I am unpopular.
18. I feel money is very important.
19. I am good in my school work.
20. I can jump well.
21. I can't seem to keep my mind on school work.
22. I have nice hair.
23. I am shy.
24. I do not like school.
25. I have poor health.
26. I do not have a good figure.
27. I have lots of energy.
28. I am not a good sport.
29. I have a brother or sister that I don't like.
30. I am good looking.
31. I am hurt by criticism.

CHILD DEVELOPMENT

32. I cannot throw a ball well.
33. I dress so people will notice me.
34. I like reading.
35. I can take things apart and put them together again.
36. I wear bright colors.
37. My parents let me decide things for myself.
38. I like my parents.
39. I am clumsy.
40. I am weak.
41. I cannot make up my mind.
42. My parents expect too much of me.
43. I get places on time.
44. I am tall.
45. I do not like active games.
46. I am afraid to take chances.
47. I have confidence in my own abilities.
48. I watch and listen to TV and radio.
49. I take part in class discussions.
50. I am an unhappy person.

The reliability of the Q-sort for its use with children was ascertained by the test-retest method using all subjects. The Q-sort was found to have a reliability coefficient of .65, which was considered acceptable for the purposes of this study. Comparing the Q-sort results with those obtained by an external measure of self-concept produced a mean percentage of agreement of 70 per cent, indicating that the Q-sort is a valid instrument for studying the self-concepts of children in groups.

A sociometric questionnaire was developed and used to obtain peer groups' perceptions of children's self-concepts. Twenty-five self-referent statements taken from the 50-item Q-sort were phrased as questions, e.g. "Which children seem to look on the bright side of things?" Children were asked to write down the names of as many as four classmates in response to each of the questions.

COLLECTION AND ANALYSIS OF DATA

The instruments of the study were administered and data were collected three times during the period of the study, November, 1954, to May, 1955. In each collection of data all children present in each of the eight classrooms completed on successive days the self-sort, the sociometric questionnaire, and the ideal sort. During each collection of data each teacher performed four Q-sorts, sorting the 50 statements as she believed four children whom she felt she knew very well would sort them.

Q-correlations were computed for the teacher's and child's sortings of the Q-sort statements. Fisher's " r to z " transformation tables were used to convert each correlation to a z score. Sociometric data were tabulated and grouped in a manner which made it possible to test hypotheses dealing with peer groups' perceptions of children's self-concepts. The 5 per cent

point of significance was designated as the point at which the null hypothesis was to be rejected.

FINDINGS

Teachers' Perceptions of Children's Self-Concepts

Q-correlation coefficients showing the correspondence between the teacher's sorting and the child's sorting of 50 self-referent statements are reported in Table 1 for each of four children in eight classrooms and for each of the three times data were collected. The mean r for all correlations of teachers' and children's placement of Q-sort items during the three times data were collected was .41. This may be regarded as a fairly high correlation, especially when one notes that the grand mean reliability coefficient of the children in this sample was only .65. Correlations of the teacher's and the child's placement of 50 Q-sort items differ significantly from zero at the 5 per cent level when $r = .279$ and at the 1 per cent level when $r = .361$. Examination of the data in Table 1 indicates that all teachers had coefficients of correspondence above the 1 per cent level for at least one child. Of the 96 performances of the Q-sort by the eight teachers, 61 resulted in correlation coefficients above .36. A sign test indicates that this result is significant at the 1 per cent level. These findings, then, tend to sustain hypothesis number one, for they show that teachers' perceptions of children's self-concepts are in general positively and significantly related to these children's expressed self-concepts.

The z scores of the Q-correlations reported in Table 1 were also tested by analysis of variance to determine whether groups of teachers who represented differences in social-emotional climate, participation in child study, teacher acceptance of self and others, and grade level differed significantly in the correspondence between their perceptions of children's self-concepts and these children's expressed self-concepts. In addition, the analysis of variance revealed whether the correspondence differed significantly for the three times data were collected or for each of the four children. F ratios and probabilities obtained from the analysis of variance are reported in Table 2.

Results of the analysis of variance show that teachers who have completed three years of child study achieve significantly greater correspondence between their perceptions of children's self-concepts and these children's expressed self-concepts than do teachers who have never participated in child study. This finding, significant at the 2 per cent level, is consistent with the hypothesized prediction favoring teachers who have participated in child study.

The analysis of variance showed, too, that sixth grade teachers and teachers who are "less" accepting of self and others register significantly greater correspondence between their perceptions of and children's expressed self-concepts than do, respectively, fourth grade teachers and

TABLE I

Q CORRELATION COEFFICIENTS AND r OF Mean z FOR TEACHER'S SELF-SORT ON CHILD VS. CHILD'S SELF-SORT

Child	Child's Reliability Coefficient	I Dec. Q	II Feb. Q	III May Q	Corresponding r of Mean z
<i>Teacher A: 4th—CS—GC—More</i>					
A 6	.42	.03	.46	.54	.36
A 19	.76	.14	-.06	.12	.07
A 29	.84	.51	.55	.62	.56
A 31	.67	.17	.04	.21	.14
Corresponding r of Mean z		.22	.27	.40	.30
<i>Teacher B: 4th—CS—TC—Less</i>					
B 2	.12	.53	.36	-.03	.30
B 9	.82	.59	.51	.07	.41
B 26	.52	.45	.47	.40	.44
B 29	.75	.39	.45	.45	.43
Corresponding r of Mean z		.49	.45	.23	.40
<i>Teacher C: 6th—CS—TC—More</i>					
C 4	.35	.52	.40	.54	.49
C 10	.74	.48	.42	.52	.47
C 26	.88	.51	.31	.40	.41
C 29	.68	.62	.46	.46	.54
Corresponding r of Mean z		.54	.40	.48	.48
<i>Teacher D: 6th—CS—GC—Less</i>					
D 3	.79	.60	.50	.70	.61
D 10	.62	.48	.65	.72	.63
D 40	.65	.53	.66	.66	.62
D 41	.91	.69	.70	.72	.70
Corresponding r of Mean z		.58	.63	.70	.64
<i>Teacher J: 4th—NCS—GC—More</i>					
J 5	.39	.24	.22	.34	.27
J 8	.42	.27	.54	.27	.37
J 13	.51	-.08	.14	.06	.04
J 31	.78	.23	.56	.66	.50
Corresponding r of Mean z		.17	.38	.35	.30
<i>Teacher K: 4th—NCS—TC—Less</i>					
K 27-31-1	.62	-.01	.16	.17	.11
K 13	.56	.57	.36	.40	.45
K 35	.76	.36	.24	.39	.33
K 29-42	.69	.48	.66	.62	.59
Corresponding r of Mean z		.36	.37	.41	.38
<i>Teacher L: 6th—NCS—GC—Less</i>					
L 3-5	.68	-.02	.07	.12	.06
L 13	.62	.52	.53	.70	.59
L 29	.90	.60	.77	.75	.71
L 33	.59	.62	.50	.42	.52
Corresponding r of Mean z		.46	.51	.54	.50
<i>Teacher M-N: 6th—NCS—TC—More</i>					
M 4-8	.38	.42	.10	.34	.29
M 5	.74	.28	.01	.36	.22
M 9	.46	.32	.15	.42	.30
M 31	.53	.17	.00	.23	.13
Corresponding r of Mean z		.30	.06	.34	.24

NOTE.—Children numbered 1-24 are boys.
 Children numbered 25-48 are girls.
 4th and 6th refer to grade level.
 CS—Has completed child study.
 NCS—Has never participated in child study.

GC—Group-centered climate.
 TC—Teacher-centered climate.
 "More" is teacher more accepting of self and others.
 "Less" is teacher less accepting of self and others.

TABLE 2

F RATIOS AND PROBABILITIES OF SIGNIFICANT DIFFERENCES FOR INDEPENDENT VARIABLES IN CORRESPONDENCE BETWEEN TEACHERS' PERCEPTIONS OF AND CHILDREN'S EXPRESSED SELF-CONCEPTS

<i>Basis of Comparison</i>	<i>F</i>	<i>p</i>
Climate	3.47	
Child Study	6.91	.02
Teacher Acceptance	12.83	.01
Grade Level	11.24	.01
Time	.64	
Children	1.98	

teachers who are "more" accepting of self and others. Both findings, significant at the 1 per cent level, are contrary to the hypothesized predictions which favored fourth grade teachers and teachers who are "more" accepting of self and others. In testing the effects of climate it was found that group-centered teachers in general do not differ significantly from teacher-centered teachers in the correspondence between their perceptions of and children's expressed self-concepts.

Thus, the findings of this study support only one part of the general hypothesis which relates to teachers' perceptions of children's self-concepts. Only teachers who participated in child study showed significantly greater correspondence between their perceptions and children's perceptions of the latter's self-concepts. For two other comparisons, grade level and teacher acceptance of self and others, results were obtained which were directly opposite to the hypothesized predictions.

In this analysis the "time" factor was not significant, indicating that teachers in general do not change significantly toward greater correspondence between their perceptions of and children's expressed self-concepts. A further analysis of the data in Table 1, however, indicates that the number of higher correlations showing greater correspondence of teachers' perceptions of and children's expressed self-concepts in May as compared with the initial correlations in December is significant at the 5 per cent level. These findings, then, support the second hypothesis with qualification. Teachers change significantly in the number of their perceptions of children's self-concepts which show a greater correspondence to these children's expressed self-concepts, but the magnitude of change toward greater correspondence is not significant. Teachers do not differ significantly for different children or for sex of child in the correspondence between their perceptions of and children's expressed self-perceptions.

Graphic representations of these results showing the mean z correspondence scores for teacher groups differing with respect to the four variables at the three points of time are shown in Figures 2 and 3. These figures

CHILD DEVELOPMENT

reflect generally the findings reported above. It is interesting to note that the correspondence achieved by all teachers between their perceptions of and children's expressed self-concepts was slightly greater in December at the beginning of the study than in February, three months later. This was markedly evident for teacher-centered teachers and to a lesser extent for "more" accepting and sixth grade teachers. Since differences between the three points of time were not significant, these observed differences could be explained as due to chance fluctuations of sample. On the other hand, they may indicate that early in an experience one's initial perceptions are more accurate and valid than are those perceptions which immediately follow. The greater mean z correspondence scores in May as compared with December and February are strikingly shown in these figures. Only

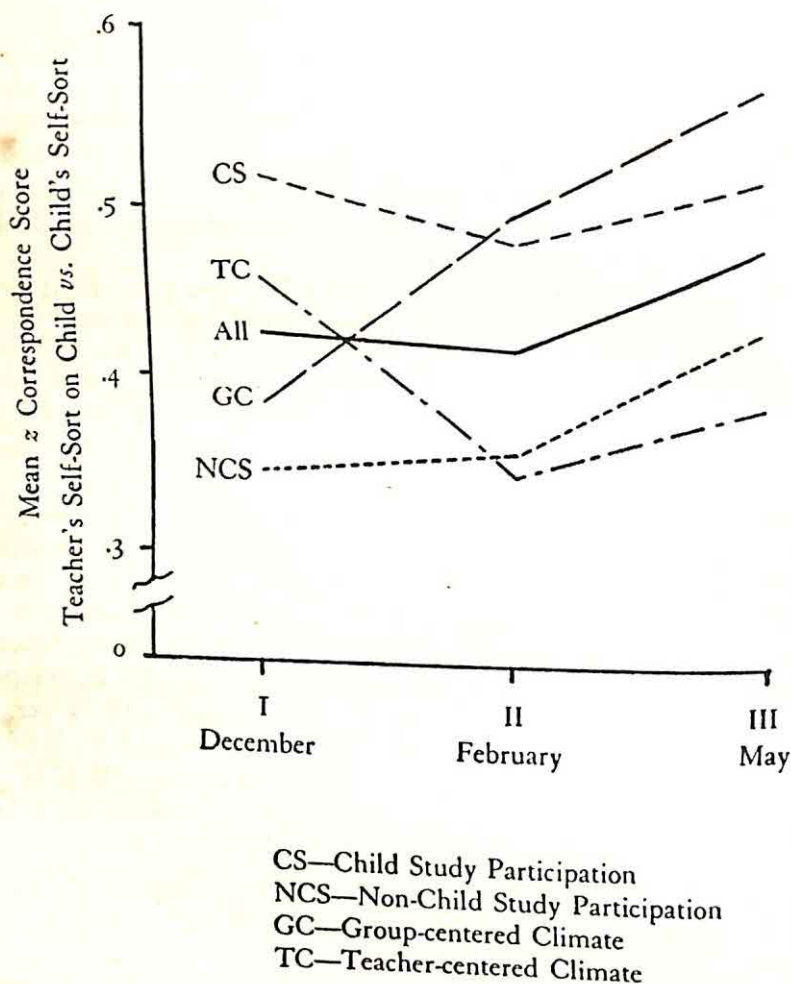
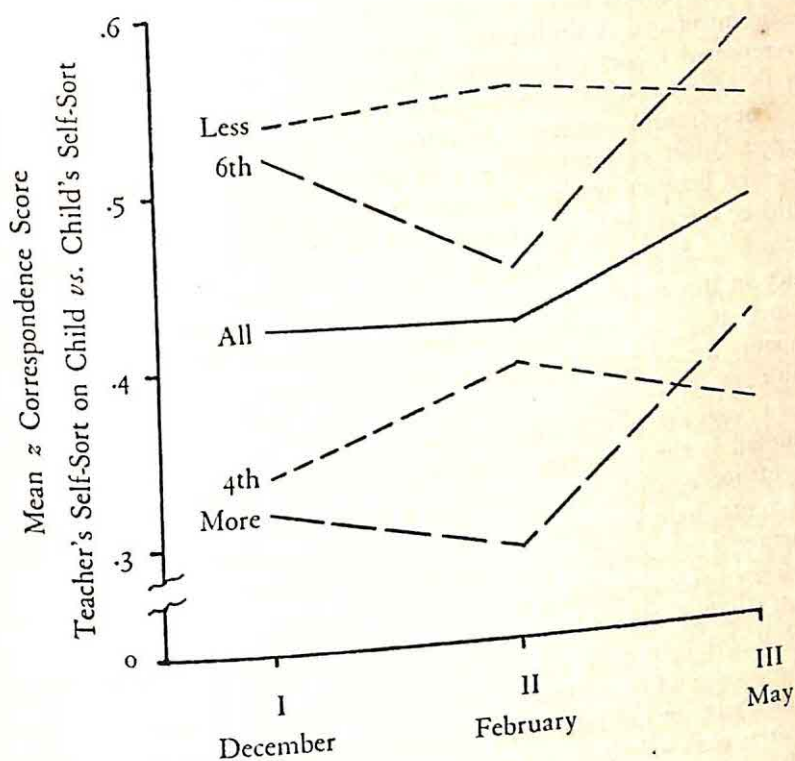


FIGURE 2—Mean z correspondence scores at three points in time for teachers who differ with respect to child study participation and social-emotional climate.

teacher-centered teachers failed to achieve a higher mean z correspondence score in May than they did at the beginning of the study in December.

A general question which is raised by these findings is: What factors might account for the observed discrepancies between teachers' perceptions of and children's expressed self-concepts? In an effort to discover some of the factors involved, Q correlation coefficients were computed for teachers and children between their first and second and their second and third performances of the Q -sort. These provided measures of consistency of the teacher's and child's perceptions of the latter's self-concept. Teachers' correlations of consistency varied from $-.04$ to $.90$ with a mean consistency coefficient of $.71$. The range of children's consistency coefficients was from $.02$ to $.86$ with a mean of $.60$.



4th—Fourth Grade Teachers

6th—Sixth Grade Teachers

More—Teachers more accepting of self and others

Less—Teachers less accepting of self and others

FIGURE 3—Mean z correspondence scores at three points in time for teachers who differ with respect to grade level and teacher acceptance of self and others.

An analysis was made to determine the relationships between consistency of teachers and children on repeated performances of the Q-sort and the correspondence between teachers' perceptions of children's self-concepts and these children's expressed self-concepts. An $r = .50$ was arbitrarily selected as the point which would distinguish "high" and "low" correspondence scores. When frequencies for total "high" and "low" correspondence scores were tested for independence with patterns of teachers' and children's consistency on repeated Q-sort performances, the chi square analysis revealed significant differences at the 5 per cent level.

In a second analysis the mean z scores and corresponding reliability coefficients of each teacher's consistency for repeated Q-sort performances were ranked, and ranks were assigned to the mean correspondence scores for the eight teachers. Rank-order correlations of ranked mean consistency scores and ranked mean correspondence scores resulted in a correlation coefficient of .50. A similar rank-order correlation of children's consistency for repeated Q-sort performances and ranked mean correspondence scores for the eight teachers produced a correlation coefficient of .60.

These results suggest that consistency of children's and teachers' repeated Q-sort performances are in general related to the correspondence achieved between teachers' perceptions of children's self-concepts and these children's expressed self-concepts. Consistency of children's repeated Q-sort performances would seem to be of greater importance than teacher consistency in this respect. The lack of high correlations between these various factors suggests, however, that other factors such as teacher insight and empathy are of greater importance in influencing teachers' perceptions of children's self-concepts.

In view of the greater correspondence which child study teachers achieved in their perceptions of and children's expressed self-concepts, one might question whether the influence of the child study experience has the effect of producing greater ability to understand an individual child or whether it is the result of a more accurate stereotyping of predictions. An answer to this question was sought by correlating the teacher's sorting of cards for Boy A and her sorting for Boy B; and likewise for Girl A and Girl B. High positive correlations of teacher's sorting for children of the same sex would be evidence of some teacher's sorting for children of the same sex. Also may be an indication of the teacher's stereotyping by sex group and children with values similar to her own. For three of the four child study teachers, Q correlations on different children of the same sex were found in general to be .50 and above, and the mean Q-correlation for all four teachers and both sexes was .46. Comparable Q-correlations for teachers who had not participated in child study were in general mixed with a mean of .31. Thus, greater correspondence in perception of self-concepts by child study teachers may reflect an increased acceptance of children due in part to stereotyping. This result can be viewed as a positive outcome of child study experience.

Peer Groups' Perceptions of Children's Self-Concepts

Data on children's perceptions of their fellow classmates' self-concepts were obtained by administering a sociometric questionnaire which asked each child to name up to four classmates whom he believed fit each of 25 qualities of self-reference which were taken from the 50-item Q-sort. The frequency "5" was selected as the "critical frequency" since for most qualities and classrooms being named five or more times by peers was significantly greater than chance at the 5 per cent point. The child's and peer group's perceptions of the child's self-concepts were tabulated as in "agreement" if the child had placed in Q-sort columns 1, 2, or 3 that item in relation to which the peer group had named him at least 5 times. If the child had placed the particular item in column 4, it was rated as "uncertain." If the child had placed the particular item in columns 5, 6, or 7 (Not Me), it was tabulated as a "disagreement." The teacher's and child's placements of like Q-sort items were also tabulated as "agreement," "uncertain," or "disagreement" in the manner described above.

A chi square test of independence comparing the frequencies of agreement, uncertainty, and disagreement for teachers' and peer groups' perceptions of the same children's self-concepts revealed no significant differences. Therefore, the results of this study do not support hypothesis 3, for they show that fourth and sixth grade teachers' perceptions of selected children's self-concepts are not significantly different from those of fourth and sixth grade peer groups in perceiving the respective children's self-concepts. It should be noted, however, that these comparisons were limited to the four children in each classroom for whom the teacher had performed the Q-sort, and they were further limited to those items for which these children had been named by their classroom peers 5 or more times.

A further analysis revealed that the perceptions of peer groups in group-centered classrooms and those whose teachers have completed child study show a significantly greater agreement with children's expressed self-concepts than do the perceptions of peer groups in teacher-centered classrooms and those whose teachers have not participated in child study. Differences in peer group perceptions relating to grade level and teacher acceptance of self and others were not significant. Thus, the general hypothesis concerned with the influence of selected factors on peer groups' perceptions of children's self-concepts is sustained in part, only as it related to climate and teacher participation in child study.

In order to test hypothesis 4 the list of 25 items on the sociometric questionnaire was arbitrarily divided into two groups: (a) a group of 11 "less threatening" percepts, and (b) a group of 14 "more threatening" percepts. Tests of significance revealed that differences of climate and teacher participation in child study were significant at the 1 per cent level for "less threatening" percepts but none of the differences relating to the four selected factors was significant for "more threatening" percepts. Thus, hypothesis 4 is in general supported. These findings suggest that differences

between the two kinds of percepts may be explained by the hypothesis that "more threatening" percepts attributed to a child by the peer group cannot be accepted by the child because they are inconsistent with the way the child wishes to see himself.

DISCUSSION

The positive and significant changes in teachers generally and in child study teachers in particular toward greater correspondence between their perceptions of children's self-concepts and these children's expressed self-concepts sustain the hypotheses with respect to these factors. The lack of significant change in magnitude of greater correspondence between teachers' perceptions of children's self-concepts and these children's expressed self-concepts may be explained by the stable and consistent quality of the teacher's self-concept. Because of this quality her ways of seeing things, including her perceptions of children, probably do not change radically during a period of time comparable to that of this study.

A finding which was clearly not anticipated is the significantly greater correspondence in perceptions of children's self-concepts which were achieved by teachers who are "less" accepting of self and others. One explanation which may account for this finding relates to the limitation of the methodology and instrument used in this study. In this respect it should be noted that differences between teachers in acceptance of self and others were far less marked than for the other three variables. Classrooms comprising the sample were actually selected in terms of differences in grade level, teacher participation in child study, and social-emotional climate. The variable of teacher acceptance of self and others was added after the beginning of the study in an effort to identify additional factors which influence teachers' perceptions of children's self-concepts. Since the scores for four teachers were bunched around the mean, it is possible that the Reed Sentence Completion Test did not differentiate clearly two teacher groups as "more" and "less" accepting of self and others.

An equally plausible hypothesis would be one which modifies present theory. It may well be, as these findings show, that persons who are "less" accepting of self and others are actually more insightful in their perceptions of others' self-concepts. Bills (1), for example, found that people who are less accepting of themselves but who believe that other people are more accepting of themselves are highly accurate in their perceptions of reality whereas persons who are self-accepting but who believe others do not accept themselves are quite inaccurate in the perceptions (p. 23). Teachers who are "less" accepting may well develop greater empathy with and insight into feelings of others because of their own heightened sensitivity in the area of interpersonal relationships. On the other hand, teachers who are "more" accepting of self and others may concentrate their energies on dealing with the problem or situation because for them accept-

ance emerges into figure less frequently. In short, they tend to accept children as they are without probing too deeply to discover how these youngsters see and feel about themselves.

The superiority of sixth grade teachers over fourth grade teachers in the correspondence between their perceptions of and children's expressed self-concepts is probably due to the greater stability and consistency of self-concept which is revealed in sixth grade children. This is suggested by the finding that a fairly high positive relationship exists between the consistency of children's repeated Q-sort performances and the correspondence achieved between teachers' perceptions of children's self-concepts and these children's expressed self-concepts. The consistency and stability of children's self-concepts, therefore, would seem to be a more significant factor influencing the correspondence achieved by teachers than is the hypothesized tendency for adults to have less insight and knowledge of children's groups with the increasing maturity of the students.

It should be noted that other factors not identified or evaluated by this study may be more significant in influencing teachers' perceptions of children's self-concepts. Differences in perception, in sensitivity to feelings of others, or in ability or inclination to empathize with others may well be of greater significance in accounting for teacher differences. The ability of a person to make valid hypotheses of another person's self-concepts may be a unique quality of the teacher's self. A teacher may have this skill to a greater or lesser degree because of her unique life experiences and the meanings and integrations she has made of them. These are factors that should be studied by further research.

The lack of significant differences between teachers and peer groups in their perceptions of children's self-concepts may be due to the small number of children (four) and limited number of percepts upon which these comparisons are based. If only those percepts were studied which are easily observed and fairly obvious to teachers and peers alike, there would be small likelihood for significant differences to emerge. This hypothesis should be tested by further research utilizing a larger sample of teachers, peer groups, and number of percepts.

The open and frank responses of fourth graders may result in fairly valid perceptions of their self-concepts by peers and others. This may more than compensate for the more structured peer group and the greater skill, understanding, and maturity which would benefit sixth graders in their perceptions of children's self-concepts.

An alternative explanation would suggest that maturity of groups may not be the most crucial factor which influences peer groups' perceptions of children's self-concepts. This explanation is supported by the findings of children's self-concepts. This explanation is supported by the findings which indicate the superiority of peer groups in group-centered classrooms and of those who are taught by child study teachers. Indeed, classrooms which have a warm, acceptant climate and those wherein teachers seek to understand children and to facilitate their growth and development

would seem to be situations which would enable children to perceive their fellow classmates' self-concepts more objectively. Similarly, greater tension in teacher-centered climates and less understanding of children's needs by non-child study teachers would tend to result in peer groups' perceptions of children's self-concepts showing less correspondence with these children's expressed self-concepts. This conclusion and interpretation seems to be warranted by the findings of this study.

Finding that peer groups' perceptions show significantly greater agreement with children's expressed self-concepts for "less threatening" percepts than for "more threatening" percepts is consistent with the hypothesized and commonly observed effects of conflict and threat on learning and behavior. Individuals are reluctant to accept unfavorable percepts into their self-concept, and a frequent way of adjusting to threatening percepts is to relegate them to the "uncertain" and "not me" aspects of self. This would seem to suggest that there is a need for education as well as therapy to provide experiences which would enable persons to become objective, open, and acceptant of all experiences, the unfavorable as well as the favorable percepts.

SUMMARY

This study investigated the influence of four factors: (a) social-emotional climate, (b) teacher participation in an in-service child study program, (c) teacher acceptance of self and others, and (d) grade level on perceptions of and changes in the self-concepts of children in the fourth and sixth grades. This article reports the findings relating to the influence of the above four factors on teachers' and peer groups' perceptions of children's self-concepts.

The general hypothesis of the study postulated that teachers and peer groups in classrooms which are group-centered, in which the teacher has participated in child study, and those where she was more accepting of self and others would reveal in their perceptions a greater correspondence with children's expressed self-concepts than would teachers and peer groups who differ with respect to these variables. In addition, greater correspondence of perceptions was predicted for fourth grade teachers and sixth grade peer groups.

A representative sample consisting of four fourth grade and four sixth grade classrooms in seven elementary schools of a suburban county school system in Maryland was selected so that extremes with respect to each of the four variables were represented. Classroom climates, group-centered and teacher-centered, were identified using Withall's Climate Index. Teachers who had completed the three-year in-service child study program sponsored by the Institute for Child Study of the University of Maryland were compared with teachers who had never participated in this program. A 50-item Q-sort was developed and used to obtain measures of all children's self-concepts, and it was also used by teachers to register their perceptions

of selected children's self-concepts. A sociometric questionnaire was developed and used to obtain peer groups' perceptions of children's self-concepts.

The findings of this study are:

1. Teachers' perceptions of children's self-concepts are in general positively and significantly related to these children's expressed self-concepts.
2. Teachers change significantly through time in the increased number of their perceptions of children's self-concepts which show a greater correspondence to those children's expressed self-concepts, but the magnitude of change is not significant.
3. Sixth grade teachers, teachers who have completed three years of child study, and teachers who are less accepting of self and others show in their perceptions of children's self-concepts a significantly greater similarity to these children's expressed self-concepts than do, respectively, fourth grade teachers, teachers who have never participated in child study, and teachers who are more accepting of self and others.
4. Neither fourth grade teachers nor sixth grade teachers show in their perceptions of children's self-concepts any significantly greater similarity with these children's expressed self-concepts than do fourth and sixth grade peer groups in their perceptions of the same children's self-concepts.
5. Peer groups in group-centered climates and those taught by teachers who have completed child study show in their perceptions of children's self-concepts a greater similarity with these children's expressed self-concepts than do peer groups, respectively in teacher-centered climates and those taught by non-child study teachers.
6. Peer group perceptions of children's self-concepts are significantly more like these children's expressed self-concepts for "less threatening" perceptions than they are for "more threatening" perceptions.

From these findings it is concluded that teacher participation in child study is the one of four factors studied which exerts the greatest influence on teachers' and peer groups' achieving increased correspondence between their perceptions of and children's expressed self-concepts. The superiority of child study teachers is a measure of the effectiveness of this in-service program which seeks to help teachers gain a deeper understanding of children's behavior. The results of this study suggests a revision of theory which may be expressed in terms of the following postulate: Teachers who are "less" accepting of self and others are more accurate and insightful in their perceptions of others' self-concepts than are teachers who are "more" accepting of self and others. The implications of this and other findings for further research are discussed.

REFERENCES

1. BILLS, R. E. About people and teaching. *Bull. Bureau Sch. Serv., Coll. Educ., Univer. of Kentucky*, 1955, 28.

CHILD DEVELOPMENT

2. CHODORKOFF, B. Adjustment and discrepancy between the perceived and ideal self. *J. clin. Psychol.*, 1954, 10, 266-268.
3. HARTLEY, MARGARET. Changes in the self-concept during psychotherapy. Unpublished doctor's dissertation, Univer. of Chicago, 1951.
4. JERSILD, A. T. *In search of self*. New York: Teachers Coll., Columbia Univer., 1952.
5. KINDRED, R. B. An investigation into the relationship of the use of human development principles to training and teaching experience. Unpublished doctor's dissertation, Univer. of Maryland, 1954.
6. LECKY, P. *Self-consistency: a theory of personality*. New York: Island Press, 1945.
7. LEWIN, K. *A dynamic theory of personality*. New York: McGraw-Hill, 1935.
8. MOUSTAKAS, C. E. *The teacher and the child*. New York: McGraw-Hill, 1956.
9. RAIMY, V. C. The self-concept as a factor in counseling and personality organization. Unpublished doctor's dissertation, Ohio State Univer., 1943.
10. REED, H. J. An investigation of the relationship between teacher effectiveness and the teacher's attitude of acceptance. *J. exp. Educ.*, 1953, 21, 277-325.
11. ROGERS, C. R. *Client-centered therapy*. Boston: Houghton-Mifflin, 1951.
12. ROGERS, C. R., & DYMOND, ROSALIND F. *Psychotherapy and personality change*. Chicago: Univer. of Chicago Press, 1954.
13. SNYGG, D., & COMBS, A. W. *Individual behavior*. New York: Harper, 1949.
14. STEPHENSON, W. *The study of behavior: Q-technique and its methodology*. Chicago: Univer. of Chicago Press, 1953.
15. WITHALL, J. G. The development of a technique for the measurement of social-emotional climate in the classroom. *J. exp. Educ.*, 1949, 17, 347-361.

FACTORS INFLUENCING CHANGE IN CHILDREN'S SELF-CONCEPTS

HUGH V. PERKINS

*Institute for Child Study
University of Maryland*

In a previous article (5) this writer suggested that the self-concept phenomenon can serve education in two ways: (a) as a psychological construct which enables teachers, counselors, parents and others to achieve with training deeper understandings and insights into the behavior and development of children, and (b) as a vital and important aspect of learning and development which the school through its educational processes seeks to promote and foster in every child. The need for a further examination and understanding of the uses of this construct in education provided the focus for an investigation made by this writer of selected factors influencing perceptions of and changes in children's self-concepts. A previous article reported the methodology and the findings of this research which related to teachers' and peer groups' perceptions of children's self-concepts (5). This article will report the findings relating to the influence of selected factors on changes in children's self-concepts.

This second part of the study investigated the effects of four factors, (a) social-emotional climate, (b) teacher participation in an in-service child study program, (c) teacher acceptance of self and others, and (d) grade level, on changes in children's self-concepts in the fourth and sixth grades.

Self-concept was defined as those perceptions, beliefs, feelings, attitudes, and values which the individual views as describing himself. The ideal self was defined in terms of those qualities which describe the person he would like to be. In this study data on children's self-concepts and ideal selves were obtained by means of a Q-sort, an instrument and technique which requires the subject to make a forced normal distribution of a series of self-referent statements. The Q-sort developed for use in this study was a stratified random sample of 50 self-referent statements which had been selected from a universe derived from responses of fourth and sixth grade children reported in a study by Jersild (3). Correlating the child's self-sort with his ideal sort provides a measure of his self-ideal self congruency and may be expressed in terms of a Q-correlation coefficient. Change in self-concept in this study was defined as change in self-ideal congruency.

A representative sample of eight classrooms enrolling 251 children in seven suburban elementary schools of a large county school system in Maryland was selected for study. Classrooms were selected on the basis of their being willing to participate in the research and their being representative of one of the extremes with respect to each of the independent variables: (a) social-emotional climate, group-centered *vs.* teacher-centered,

CHILD DEVELOPMENT

(b) participation in child study, completion of the three year program in child study *vs.* no participation in program, (c) teacher acceptance of self and others, more acceptance *vs.* less acceptance of self and others, and (d) grade level, fourth *vs.* sixth.

Classroom climates, group-centered and teacher centered, were identified using Withall's Climate Index (7). Participation in child study referred to teacher participation for three years in the in-service child study program sponsored by the Institute for Child Study of the University of Maryland. Teacher scores on Reed's Sentence Completion Test (6) were used as measures of teacher acceptance of self and others. Scores on the California Test of Mental Maturity, 1953 S Form, and the California Achievement Tests in Language, Reading, and Arithmetic, Primary and Intermediate Forms A and B, were used in ascertaining the relationships between reliability in using the Q-sort and reading and mental age and in correlating changes in school achievement with changes in self-ideal self congruency.

The general hypothesis tested by the second part of this study was as follows: Children in group-centered climates, those whose teachers have participated in child study, those whose teachers have a greater acceptance of self and others, and those in sixth grade will show greater congruency between their self-concepts and ideal selves than will children, respectively, in teacher-centered climates, those whose teachers have never participated in child study, those whose teachers have less acceptance of self and others, and those in fourth grade.

In addition, the following specific hypotheses were tested:

1. Children generally will register positive and significant changes through time toward increased congruency between their self-concepts and ideal selves.
2. Girls will register a significantly greater congruency between their self-concepts and ideal selves than will boys.
3. Children who register greater change in school achievement and peer group acceptance will also show greater change toward congruency between their self-concepts and ideal selves.

FINDINGS

Reliability of Q-sort Instrument

The results of the reliability test of children in performing the Q-sort, though not a major problem of this investigation, are reported because of the clues they suggest regarding the stability and consistency of children's self-concepts. This test of reliability of children's use of the Q-sort was accomplished by asking the children in each classroom to make a repeated self-sort (self-concept) two to seven days after they had made an initial self-sort. This test of reliability was made during the second collection of data after the children had become familiar with the instruments. Mean z scores and corresponding reliability coefficients for repeated administration of the Q-sorts are reported in Table 1 by sex, classroom, and grade.

HUGH V. PERKINS

The reliability coefficient of .65 for the Q-sort, though somewhat lower than had been anticipated, indicates, nevertheless, that the Q-sort is a reliable instrument for measuring the self-concepts of children in groups. Since the reliability test of an instrument is also a test of the reliability of persons using the instrument, the observed reliability coefficients of fourth and sixth grade children's use of the Q-sort in this study is indica-

TABLE I

MEAN z SCORES AND CORRESPONDING RELIABILITY COEFFICIENTS
OBTAINED FROM REPEATED ADMINISTRATIONS OF Q-SORT

Teacher and Class	B O Y S			G I R L S			ENTIRE CLASS	
	N	Mean z	r	N	Mean z	r	N	Mean Corresponding r
<i>Fourth Grade:</i>								
A	18	.616	.55	20	.736	.63	38	.679
B	9	.801	.66	6	.766	.64	15	.787
J	18	.418	.40	18	.785	.66	36	.601
K	16	.651	.51	18	.760	.64	34	.666
Entire Fourth	61	.599	.54	62	.762	.64	123	.683
<i>Sixth Grade:</i>								
C	13	.802	.66	18	.924	.73	31	.876
D	20	.816	.67	18	1.028	.77	38	.916
L	13	.970	.75	10	.935	.73	23	.955
M	13	.674	.59	23	.810	.67	36	.761
Entire Sixth	59	.816	.67	69	.924	.73	128	.877
Entire Sample	120	.707	.61	131	.843	.69	251	.780

tive of the level of consistency of their self-concepts at these maturity levels. Thus, the relative consistency of children's self-concepts related not only to the reliability of the Q-sort instrument, but it was also an important aspect of the phenomenon studied by the research.

Other results of the test for reliability reported in Table I indicate that girls' reliability in using the Q-sort instrument was .69 whereas the boys' was .61. Moreover, sixth grade children, whose coefficient was .70, appeared to be more reliable than fourth graders whose coefficient was .59. These differences in reliability between boys and girls and between fourth and sixth grade children were analyzed for significant differences using the normal curve sampling distribution. For a two-tailed test of significance, neither the z value of 1.06 for sex differences nor the z value of 1.51 for grade level differences was significant. Since the z value of 1.51 approaches significance at the 5 per cent point for a one-tailed test, the hypothesis that more mature children revealed greater consistency in their self-concepts than do less mature children should be tested by further research.

CHILD DEVELOPMENT

In this research it was assumed that children in fourth and sixth grades have sufficient insight into their selves and sufficient skills in reading and understanding to enable them to use the Q-sort instrument. This assumption suggests that children's reliability in using the Q-sort may be related to reading level and mental age. To ascertain whether such relationships exist, product-moment correlation coefficients were obtained by correlating z scores of reliability coefficients with grade level scores in reading and with mental age. These are reported in Table 2.

TABLE 2
PRODUCT-MOMENT CORRELATION COEFFICIENTS BY CLASSROOMS
OBTAINED BY CORRELATING Q-SORT RELIABILITY z SCORES
WITH READING LEVELS AND MENTAL AGES

<i>Teacher & Classroom</i>	<i>Grade</i>	<i>Q-SORT RELIABILITY z SCORE vs.</i>	
		<i>Reading Level</i>	<i>Mental Age</i>
A	4	.57	.52
B	4	.12	.43
J	4	.64	.36
K	4	.14	.08
C	6	.26	.27
D	6	.51	.51
L	6	.20	.08
M	6	.46	.43
Corresponding r of Mean z		.38	.34

The results reported in Table 2 indicate generally that relationships between Q-sort reliability and reading level and mental age are positive but low. These findings are corroborated by those of Brandt (1) who found a low positive relationship between accuracy of self-estimate and intelligence. In the present study it will be noted that correlations of Q-sort reliability with reading level and mental age are quite similar for most classrooms. This suggests that reading performance correlates highly with mental age since Q-sort reliability operated as a constant in these comparisons. From these results it is concluded that other factors such as personal adjustment are probably more closely related to consistency of self-concept than are reading level and mental age.

Changes in Children's Self-Concepts

Change in self-concept has been defined as change in congruency of the child's self-concept and his ideal self occurring during a specified period of time. An individual may be presumed to be moving toward better adjustment and to be facilitating his self-development when his self-concept and ideal self becomes more congruent. Evidences of children's self-concepts

HUGH V. PERKINS

and ideal selves were obtained by means of the Q-sort which was administered three times during the study, in December, February, and May, 1954-1955. Q-correlation coefficients of congruency of self-concept and ideal self were recorded for each child and transformed into z scores. Mean z scores and corresponding r correlation coefficients were computed for each

TABLE 3
MEAN z SCORES AND r CORRELATION COEFFICIENTS OF CONGRUENCY
OF CHILDREN'S SELF-CONCEPTS AND IDEAL SELVES

Teacher and Class	N	Sex	I DEC.		II FEB.		III MAY		Mean z and Corresponding r	
			z	r	z	r	z	r	z	r
A 19 Boys			.382	.36	.537	.49	.618	.55	.512	.47
A 19 Girls			.418	.40	.675	.59	.633	.56	.575	.52
Mean z & r . . .			.400	.38	.606	.54	.625	.55	.544	.50
B 10 Boys			.451	.42	.485	.45	.503	.46	.480	.45
B 5 Girls			.691	.60	.835	.68	.829	.68	.785	.67
Mean z & r . . .			.571	.51	.660	.58	.666	.58	.632	.56
C 13 Boys			.604	.54	.565	.51	.662	.58	.610	.54
C 14 Girls			.578	.52	.606	.54	.688	.60	.624	.55
Mean z & r . . .			.591	.53	.585	.53	.675	.59	.617	.55
D 23 Boys			.526	.48	.625	.55	.688	.60	.613	.55
D 18 Girls			.641	.56	.815	.67	.827	.68	.761	.64
Mean z & r . . .			.584	.53	.720	.62	.758	.64	.687	.60
J 15 Boys			.420	.40	.405	.38	.382	.36	.402	.38
J 16 Girls			.545	.50	.618	.55	.679	.59	.614	.55
Mean z & r . . .			.482	.45	.512	.47	.530	.48	.508	.47
K 13 Boys			.360	.35	.355	.34	.303	.29	.339	.33
K 18 Girls			.554	.50	.493	.46	.498	.46	.515	.47
Mean z & r . . .			.457	.43	.424	.40	.400	.38	.427	.40
L 13 Boys			.639	.56	.705	.61	.770	.65	.705	.61
L 10 Girls			.432	.41	.657	.58	.688	.60	.592	.53
Mean z & r . . .			.536	.49	.681	.59	.729	.62	.648	.57
M-N . . 12 Boys			.428	.40	.399	.38	.582	.52	.470	.44
M-N . . 21 Girls			.475	.44	.481	.45	.666	.58	.541	.49
Mean z & r . . .			.452	.42	.440	.41	.624	.55	.505	.47

CHILD DEVELOPMENT

sex group, each classroom, and for each of the three times data were collected, and are reported in Table 3.

Analysis of variance tests of significance were made using the data in Table 3 to determine whether differences in congruency of children's self-concepts and ideal selves were significant for each of the four independent variables and for sex group and time. The results of these tests of significance are given in Table 4.

The results of the analysis of variance reported in Table 4 show that differences in congruency of children's self-concepts and ideal selves due to differences in sex are significant at the 1 per cent level. Differences in congruency at different points in time are significant at the 5 per cent level. These findings sustain hypotheses 1 and 2, for they show that girls generally register significantly greater self-ideal self congruency than do boys, but all children reveal significant changes toward greater congruency through time, December to May.

TABLE 4
F RATIOS AND PROBABILITIES OF SIGNIFICANT DIFFERENCES FOR SEX,
TIME, AND FOUR INDEPENDENT VARIABLES ON CONGRUENCY OF
CHILDREN'S SELF-CONCEPTS AND IDEAL SELVES

<i>Basis of Comparison</i>	<i>F</i>	<i>p</i>
Sex	9.63	.01
Time	3.71	.05
Climate	2.12	
Child Study	8.65	.01
Teacher Acceptance	2.71	
Grade Level	7.52	.01

In Table 4 the report of the tests of significant differences relating to the four independent variables shows that differences in children's self-ideal self congruency are significant at the 1 per cent level for the variables of teacher participation in child study and grade level. Differences due to climate and teacher acceptance of self and others are not significant. Sixth grade children and those taught by teachers who have completed child study, then, reveal significantly greater self-ideal self congruency than do fourth grade children and those taught by non-child study teachers, respectively.

A further analysis was made to determine the influence of the independent variables on the proportionate number of increases and decreases in children's self-ideal self congruency. This was accomplished by tabulating for each child in each classroom the frequencies of increase and decrease in self-ideal self congruency as shown by a comparison of z scores for three periods of time, December-February, February-May, and December-May. Chi square tests of independence revealed that the pro-

portionate numbers of increases and decreases in children's self-ideal self congruencies produced results similar to those obtained by the analysis of variance treatments in Table 4. The chi square results were less clear cut, however, because the effects of magnitude of change were not included. These findings, then, support the general hypothesis only with respect to sixth grades and classrooms whose teachers have completed child study in which children reveal significantly greater change in self-ideal self congruency.

This research also sought to ascertain the relationships that exist between changes in children's self-ideal self congruencies and changes in children's school achievement as shown by differences in scores made on the California Achievement Tests which were administered to the four sixth grade classrooms before and after the period of this study. Evidences of changes in children's self-ideal self congruency were obtained by computing for each child the z score difference between his z score self-ideal self congruency in December and his z score self-ideal self congruency in May. Mean product-moment correlation coefficients for the four sixth grade classrooms for relationships between changes in children's self-ideal self congruency and changes in academic achievement were .09 for language, -.06 for reading, .16 for arithmetic, and .08 for total achievement. Since these coefficients do not differ significantly from zero, it is concluded that there are no significant relationships between changes in children's self-ideal self congruencies and changes in their school achievement.

Finally, this research sought to determine the relationship that exists between changes in children's self-ideal self congruency and changes in their acceptance by the peer group. Measures of change in self-ideal self congruency used in studying this relationship were the same differences in z score congruency, December and May, which were used previously in correlating this factor with school achievement. Measures of change in peer group acceptance were obtained from the answers children made to a twenty-sixth question on the sociometric questionnaire: Who are your very best friends in this class? Change in peer acceptance was defined as the difference between the number of times a child was named as a friend in December and the number of times he was named in May. Mean product-moment correlations for the relationships of changes in children's self-ideal self congruency and changes in their peer acceptance were .04 for fourth grades and .03 for sixth grades. Mean correlations for the relationship of z score self-ideal self congruency and peer group acceptance at the end of the study in May, 1955, were .17 for fourth grades and .27 for sixth grades. These findings show that there is a lack of relationship between changes in children's self-ideal self congruency and changes in their acceptance by peers and that children with greater self-ideal self congruency are not chosen more frequently by their peers. Thus, the lack of relationships between change in children's self-ideal self congruency and changes in their school achievement and peer group acceptance indicates a lack of support in this study for hypothesis 3.

CHILD DEVELOPMENT

DISCUSSION

In this research it was posited that increase in self-ideal self congruency facilitates self-development and is evidence of the improved self-adjustment of the individual. Since a major purpose of education is to promote individual self-development, the challenge of helping individuals increase their self-ideal self congruency would seem to be of crucial importance to education.

It is reassuring to note, therefore, that according to the findings of this study children in general do achieve increased self-ideal self congruency during a period of time that they are in attendance at school. Two explanations for this increase are suggested. First, it may be influenced by the general pattern of children's growth and development. It seems reasonable to postulate that as children mature, their self-concepts become more stable and consistent and their self-concepts and ideal selves increase in congruency. This would not deny, of course, the importance of specific factors in the environment or life experience of the individual which would be likely to influence his self-ideal self congruency irrespective of the maturational factor which has been postulated. A second explanation is that the schools may well be a principal factor influencing this general increase in self-ideal self congruency since school experience is the only common factor in the lives of these children. The extent to which the increase can be attributed to school experience may be an indication of the school's effectiveness in facilitating this important aspect of children's development.

It is apparent, however, that the school is not uniformly successful in facilitating greater self-ideal self congruency in all children. The findings of this study show that boys in general achieve significantly less congruency of self-concept and ideal self than do girls. One might hypothesize that the culture generally is responsible for producing these differential changes in boys and girls, but this seems unlikely. The greater physical maturity of the average girl compared with the average boy may be a partial explanation for the observed differences which favor girls. Girls generally are ahead of boys in physical maturity until the end of pubescence when boys finally catch up to the girls. It may well be, then, that the greater self-ideal self congruency registered by girls is an interrelated aspect and concomitant of their increased physical maturity. Brandt (1) found a similar sex difference in his study. He found that boys register greater accuracy of self-estimate in achievement areas than do girls, but the tendency to overrate rather than to underrate themselves was significantly greater among boys.

It seems plausible to hypothesize, too, that in elementary school classrooms boys in general have fewer opportunities for experiences which would enable them to achieve a measure of self-ideal self congruency comparable to that achieved by girls especially when feminine cultural standards predominate in a classroom taught by a woman teacher. The sex of the teacher was controlled in this study by selecting only classrooms

taught by women teachers. If the patterns of behavior and activities in the school are largely governed by feminine mores and codes, it seems likely that boys would inevitably be less able to behave in ways which permit their achieving self-ideal self congruencies comparable to those of girls. This is an intriguing hypothesis which should be tested by further research.

The findings which show that sixth graders and children of child study teachers achieve greater self-ideal self congruency than do fourth graders and children of non-child study teachers, respectively, are especially noteworthy. The greater congruency of sixth graders in contrast to fourth graders may well be due to their greater maturity and experience. This finding parallels an earlier one which shows that the self-concepts of sixth graders tend to be more stable and consistent than are those of fourth graders. This difference relating to maturity was corroborated by Brandt (1) who found in his study that eleventh grade students were significantly more accurate in self ratings than were sixth grade students.

The greater self-ideal self congruency of children taught by child study teachers would seem to be a consequence of these teachers' having a greater sensitivity to and providing for children's developmental needs as the result of the knowledge and understanding which they had gained in their child study experience. If teachers' experiences in a program employing the direct study of children have been meaningful, it is plausible to expect that the results of child study will be reflected in the freedom and opportunities which result in children achieving greater congruency of self-concepts and ideal selves.

The lack of relationship between changes in children's self-ideal self congruencies and changes in their school achievement and acceptance by peers would seem to indicate that such changes are relatively independent of these other aspects of self-development. Since academic aspects of school life seem to have little relationship with changes in self-ideal self congruency, children who may not achieve well in academic skills can and do achieve in other areas of self-development. Finally, it is recognized that many factors besides the four studied by this research influence changes in children's self-ideal self congruency. These should be identified and studied by further research.

SUMMARY

A study was made of the influence of four factors, (a) social-emotional climate, (b) teacher participation in an in-service child study program, (c) teacher acceptance of self and others, and (d) grade level, on perceptions of changes in children's self-concepts. This article reports the findings relating to the influence of these factors on changes in children's self-concepts.

A representative sample consisting of four fourth grade and four sixth grade classrooms enrolling 251 children in seven elementary schools of a suburban county school system in Maryland was selected so that extremes

with respect to the four variables were represented. Evidences of children's self-concepts and ideal selves were obtained by having all children perform a self-sort and ideal sort three different times during a six-months period. Correlating the child's self-sort with his ideal sort provided a measure of self-ideal self congruency. Change in self-concept was in terms of change in self-ideal self congruency.

The findings of the study relative to changes in children's self-concepts are:

1. The self-concepts and ideal selves of children become increasingly and significantly congruent through time.
2. The self-ideal self congruencies of girls generally are significantly greater than those of boys.
3. Sixth grade children and children whose teachers have completed child study show significantly greater self-ideal self congruency than do children, respectively, in fourth grade and those whose teachers have never participated in this child study program.
4. There is little or no relationship between changes in children's self-ideal self congruency and (a) changes in their school achievement, and (b) changes in their acceptance by peers.

In the discussion of the findings it was hypothesized that changes in self-ideal self congruency of children in general may be a manifestation of growth and development and may reflect the positive influence of school experience. The greater self-ideal self congruencies of girls contrasted with boys may relate to their greater physical maturity and may be influenced by the school's feminine mores and codes which tend to favor girls. The greater self-ideal self congruencies of children whose teachers have participated in child study is evidence that this program benefits not only teachers but children as well.

REFERENCES

1. BRANDT, R. M. Accuracy of self-estimate: a measure of self-concept reality. *Genet. Psychol. Monogr.*, in press.
2. CHODORKOFF, B. Adjustment and discrepancy between the perceived and ideal self. *J. clin. Psychol.*, 1954, 10, 266-268.
3. JERSILD, A. T. *In search of self*. New York: Teachers Coll., Columbia Univer., 1952.
4. LEPINE, L. T., & CHODOROFF, B. Goal setting behavior, expressed feelings of adequacy and the correspondence between perceived and ideal self. *J. clin. Psychol.*, 1955, 11, 395-397.
5. PERKINS, H. V. Teachers' and peers' perceptions of children's self-concepts. *Child Developm.*, 1958, 29, 203-220.
6. REED, H. J. An investigation of the relationship between teacher effectiveness and the teacher's attitude of acceptance. *J. exp. Educ.*, 1953, 21, 277-325.
7. WITHALL, J. G. The development of a technique for the measurement of social-emotional climate in the classroom. *J. exp. Educ.*, 1949, 17, 347-361.

THE ROLE OF INCENTIVE IN DISCRIMINATION LEARNING IN CHILDREN^{1,2,3}

GLENN TERRELL, JR.

University of Colorado

Many of the experiments involving the effects of different incentives on children's learning have employed relatively complex tasks (1, 3). In order to provide for more rigorous experimental control, more incentive studies involving children's learning of simple concepts need to be done. Then, as we improve our experimental methods, we can more satisfactorily study the effects of incentives on the learning of more complex tasks. Accordingly, the experiment reported here is the second in a series designed to study the effects of various incentives on simple discrimination learning and transposition in children. In a previous study (5), the differential effects of five incentives on the acquisition and transposition of a button pushing response to the larger of two three-dimensional geometric objects were studied. The five incentives employed were praise, immediate candy, reproof, delayed candy, and light flash. In this earlier problem, the Ss, pre- and elementary school children, assigned to the immediate candy incentive condition learned the concept "larger than" significantly more quickly than did the Ss of all other groups, and transposed more consistently than did any of the other groups except the delayed incentive Ss. The nonsignificant difference in performance of the immediate and delayed groups on the transposition test was somewhat surprising. Even more surprising, perhaps, was the fact that, under the delayed incentive condition, the younger Ss, four- and five-year-olds, learned just as quickly and transposed just as consistently as the older Ss, eight- and nine-year-olds. This is contrary to reports of casual observation made by teachers and other supervisors of children's learning. It was believed that the nature of the delayed reward facilitated learning and transfer. The Ss transferred a dried bean from one jar to another after each correct response, and were told that as soon as they had transferred enough of the beans they could trade

¹ This investigation was supported by a research grant, M-1505, from the National Institute of Mental Health, of the National Institutes of Health, U. S. Public Health Service.

² The writer wishes to express his appreciation to Mr. Jack W. Fleming, Graduate Research Assistant, who collected the data of the study. Grateful acknowledgement is also due the many teachers of the Martin Park Elementary School, Fairview District, Boulder, Colorado, who so generously provided the subjects of the experiment. Special thanks go to Superintendent Nevin J. Platt and Principal Melvin Wiesley, without whose cooperation this experiment could not have been performed.

³ This investigation was conducted under the auspices of the Behavior Research Laboratory, Psychology Department, University of Colorado.

them in for a small sack of candy. Thus they were able to see objective evidence of progress toward the delayed goal. It was further thought that a delayed incentive consisting of a mere promise of a sack of candy for learning the task would be much less effective than one where the *S* engages in some activity which provides evidence of progress toward the goal. The present experiment was designed to test the latter belief. Specifically, the purpose of the present experiment was to compare two types of delayed incentives with each other and with an immediate incentive in the acquisition and transposition of a button-pushing response to the larger of two three-dimensional geometric objects.

METHOD

Subjects

There were 20 *Ss* in each of the following age categories: four-, five-, eight- and nine-year-olds, with an equal number of boys and girls in each age group.

Materials

The apparatus for the problem has been described in detail elsewhere (5). There were three pairs of three-dimensional geometric figures in the shape of cubes, cones, and cylinders. The small member of each stimulus set had a basal area of four square inches, while the large member had a basal area of eight square inches. These stimuli are hereafter referred to as the training stimulus sets. A third cube with a basal area of sixteen square inches was used in a transposition test, along with the eight square inch cube. These stimuli are hereafter referred to as the test stimuli. The order of presenting the stimuli and the position of the positive, large-size stimulus were randomized alike for each *S*.

Additional apparatus consisted of a 16 by 24 by 4 inch box which contained the batteries and circuits necessary to operate a signal light. Two jacks and two push-button mounts were on top of the box. The stimuli were placed into the jacks on each trial. Locked onto the rear edge of the box was a 10 by 16 by $\frac{1}{4}$ inch panel board which contained the signal light. The circuits were arranged so that a correct response, pushing the button at the base of the large stimulus, caused the light to go on.

Design

There were four experimental groups, all of which were rewarded with a light flash after each correct response. The differential incentive conditions for each group were as follows: The *Ss* of group I, the immediate incentive group, were given a small piece of candy after each correct response. This group is hereafter referred to as the immediate group. The *Ss* of group II, the first delayed incentive group, were instructed to transfer a dried bean from one jar to another after each correct response. These *Ss* also were told that they could trade their beans in for a sack of candy

when they had earned enough of them. This group is hereafter referred to as the token group. Group III Ss, the second delayed incentive group, were simply promised a small sack of candy after they had made the light go on enough times. This group is hereafter referred to as the promise group. Group IV, the control group, received only a light flash indicating a correct response.

The design consisted of two levels of chronological age in addition to the four treatment groups. The four- and five-year-olds constituted one level; while the eight- and nine-year-olds were the other level. Within each level the Ss were randomly assigned to the four treatments, making a total of 10 Ss in each treatment-level combination.

Procedure

The Ss were tested individually. Each S received the following instructions: "This is a game where I want you to try to choose one of these (E points to the training stimulus sets), and push the button in front of the one you choose. If you are right, this little light will go on. If you are wrong, the light will not go on. Now remember, the game is to see how quickly you can learn to choose the one that makes the light go on." The last sentence was repeated after every tenth trial. Immediately after reaching a criterion of nine of 10 correct responses, each S was given a four-trial transposition test on the test stimuli. The same differential incentive conditions employed during the acquisition trials were continued during the test trials.

RESULTS

An analysis of variance was performed on each of the two criterion measures: number of trials to the criterion, and number of correct responses on the transposition test. Table 1 contains both the training and transposition data. All tests performed on the transposition data were non-significant. For the acquisition data, the tests of interaction and main effects of chronological age were nonsignificant, while the treatment differences were significant beyond the 1 per cent level. The summary of the analysis of variance of the acquisition data is in Table 2. The promise group required significantly more trials to learn the concept than did all of the other groups. All other differences were of a chance nature.

Since the Bartlett test of homogeneity of variance was significant, and since some of the within cells distributions were somewhat skewed, it seemed advisable to submit the data to a statistical test requiring no assumptions of normality nor homogeneity. For this purpose the Friedman two-way analysis of variance by ranks technique (4) was employed and the resulting χ^2 was significant well beyond the 1 per cent level. The Ss who were given a promise of a delayed incentive required more trials to learn the task at all age levels with the exception of the nine-year-level, where they were second only to the immediately rewarded group.

CHILD DEVELOPMENT

TABLE 1

MEANS AND STANDARD DEVIATIONS OF TRIALS TO CRITERION IN
TRAINING AND CORRECT RESPONSES IN TRANSPOSITION TEST(Each treatment group $N = 20$)

Group	TRAINING		TRANSPOSITION	
	Mean	SD	Mean	SD
<i>Reward:</i>				
Immediate	9.15	13.58	4.00	0.00
Token	6.15	6.59	3.80	0.87
Promise	20.08	21.05	3.75	0.62
Control	7.35	6.99	3.75	0.70
<i>Age level:</i>				
4 and 5	10.62	16.75	3.85	0.66
8 and 9	10.72	11.88	3.80	0.64

TABLE 2

ANALYSIS OF VARIANCE OF NUMBER OF TRIALS TO CRITERION

Source	df	Mean Square	F
Rewards	3	811.65	4.25**
Age levels	1	0.20	
Interaction	3	232.40	1.21
Within cells	72	190.62	

** $p < .01$.

DISCUSSION

The relationship between the immediate, token, and promise groups of this experiment was about as predicted from the results of the previously mentioned experiment (5). That is, it was expected that the Ss of the token group would learn more quickly than those of the promise group. Also, the fact that the token group learned as quickly as the immediate group is no surprise, in view of the fact that in the previous study there was no difference between these two groups on the transposition test. Both of these experiments, then, suggest that this type of delayed incentive, namely the token type, applied to this kind of learning task is just as effective as an immediate incentive.

The difference in behavior of the control group Ss of the previous and present experiments requires some consideration. In the former, the control group Ss required significantly more trials to learn the concept than did any of the other groups, while in the present study the control Ss learned as quickly as the token and immediate Ss. In the earlier study, the mean

number of trials to the criterion for the immediate, token, and control groups respectively were 14.66, 26.53, and 35.47. In the two experiments these groups were treated identically, and the same age levels were used. However, different experimenters and populations were employed. It is felt that differences in population characteristics account for the difference between the control groups of these two studies. In the former experiment, the Ss were from Tallahassee, Florida, while in the present study they were from Boulder, Colorado. The principal difference in the samples selected from these two populations was in socioeconomic level. The ratio of rural to urban children in the Florida sample was much greater than in the Colorado sample. The Ss used in the Colorado sample were from a residential area consisting almost entirely of people from business and higher professional occupations.⁴ To test the hypothesis that the socioeconomic factor is relevant here, another similar experiment was performed on a predominantly rural sample from the Boulder, Colorado area. The results of this experiment support the hypothesis, since these control group Ss behaved similarly, in relation to the other reward groups, to those of the Florida sample.

There is empirical evidence to suggest that the crucial variable distinguishing the socioeconomic groups of the previous and present experiments is the matter of intrinsic motivation (2). That is, to the Ss of the Colorado sample, good performance, irrespective of incentives offered them, was immensely important. Some support of this hypothesis is found in the fact that when the Ss of the present study were questioned after the experiment about "the best way to get them to do something," a great majority of them answered "ask me" in preference to either "promise me something" or "give me a piece of candy." Also, almost all Ss indicated that they would rather do something for the "fun of it" than be "promised something for doing it." Unfortunately, no information of this kind is available on the Ss of the previous study. This same hypothesis might be offered as a partial explanation for the fact that the immediate incentive group of the earlier experiment learned more quickly than the token group, while there was no difference between these two groups in the present study.

It appears very important to attempt some explanation of the inhibiting effect of a promised incentive, found in the present experiment. Two operational hypotheses will be offered. One is that the child's history with respect to his promises having been fulfilled or unfulfilled is relevant. Research designed to manipulate this variable experimentally by fulfilling promises to a random selection of Ss and nonfulfilling promises to another random selection is now under way. The prediction is that those Ss whose promises have been fulfilled will learn more quickly under a promise incentive than those whose promises have been unfulfilled.

⁴ That the Colorado sample was from a higher socioeconomic level than the Florida group may also account for the fact that the former Ss generally learned the task much more quickly.

Another possible explanation for the relatively slow learning of the promise group lies in the meaninglessness of a promise incentive in the acquisition of a larger-than concept. The plausibility of this hypothesis is strengthened somewhat by the fact that, as has been mentioned above, almost all of the Ss of the present study indicated that they would rather do something for the fun of it than to be promised something for it. This hypothesis might be tested by making the promise condition fun or meaningful in relation to the task to be learned. One possible way to do this would be to instruct one group of promise incentive Ss to pretend that they are filling their imaginary sack with candy that the E pretends to place before them after each correct response. The performance of a group of Ss so treated may then be compared with another group, treated as those promise Ss were in the present experiment. Such an experiment is currently in progress.

Some mention should be made of the fact that there were no treatment differences in mean number of transposition responses. Since all but seven Ss transposed on every trial, and these Ss were quite evenly distributed among the treatment groups, it appears that the simplicity of the task for these Ss masked any incentive effect on the transposition test.

SUMMARY

An experiment has been described showing the effectiveness of one type of delayed incentive, token, as compared with another, promise, and an immediate incentive, in the learning and transposition of a simple button-pushing response to the larger of two three-dimensional geometric objects. Operational hypotheses to explain the differences have been briefly described. It is believed that a systematic study of motivational variables underlying children's performance in simple tasks of this nature will greatly facilitate the eventual study of the role of incentives in the acquisition of more complex responses.

REFERENCES

1. ABEL, L. B. The effects of shifts in motivation upon the learning of a sensori-motor task. *Arch. Psychol.*, N.Y., 1936, No. 205.
2. DOUVAN, E. Social status and success striving. *J. abnorm. soc. Psychol.*, 1956, 52, 219-223.
3. HURLOCK, E. B. The value of praise and reproof as incentives for children. *Arch. Psychol.*, N.Y., 1924, No. 71.
4. SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
5. TERRELL, G., & KENNEDY, W. A. Discrimination learning and transposition in children as a function of the nature of the reward. *J. exp. Psychol.*, 1957, 53, 257-260.

PERCEPTUAL RIGIDITY AND CLOSURE PHENOMENON AS A FUNCTION OF MANIFEST ANXIETY IN CHILDREN

CHARLES D. SMOCK¹
Department of Psychiatry
University of Pennsylvania

An increasing number of investigations have been concerned with the psychological and behavioral correlates of the Taylor manifest anxiety scale (19). The Taylor scale (MAS) is coordinated with the theoretical drive (D) variable in Hull's behavior theory; i.e., scores on the test presumably reflect variations in the emotional reactivity (drive) of the individual. Predictions generated from the Hullian assumptions regarding the functional relationship between drive level and performance have been, in general, confirmed in a variety of simple and complex learning situations (e.g., 5, 7). However, the failure of several studies (1, 9, 12) to yield data consistent with these assumptions indicates the need for continued experimental efforts designed to clarify the functional properties of the manifest anxiety scale. The present investigation was designed to explore certain perceptual consequences of manifest anxiety in children.

In a recent review of experimental studies concerned with the relationship between motivation and learning, Bruner *et al.* (3) indicate that a reduction in behavioral adaptivity is often associated with high levels of "drive." These writers suggest that "... conditions requiring increased speed and efficiency of goal attainment ... (e.g., deprivation) ... may have the effect of narrowing the range of environmental cues to which the individual responds" (3, p. 9). Similar "perceptual constriction" effects have been noted to be associated with induced anxiety (18), motivational relevance and context of the task stimuli (8), and level of anxiety (13). However, the theoretical and experimental problems associated with specifying the processes which mediate the "narrowing of the perceptual field" have received relatively little attention.

One plausible explanation for the reduction in responsiveness to environmental cues is that relatively high motivational levels tend to increase

¹ The data for this study were collected while the author was a member of the staff of the Iowa Child Welfare Research Station, State University of Iowa.

perceptual rigidity and closure tendencies. Empirical support for this assumption may be found in numerous investigations of the effects of anxiety or stress on perceptual-cognitive processes. Most directly relevant here are the studies reported by Moffit and Stagner (10) and Smock (14, 15, 16) on the relationship between anxiety and performance in a variety of perceptual task situations. Data obtained in the latter series of investigations, for example, indicate that experimentally-induced anxiety increases the tendency for individuals to adhere to pre-recognition responses, to respond on the basis of minimal "information," and to be relatively retarded in their recognition of incongruous stimulus configurations. Moffit and Stagner (10), using the Taylor scale as the index of anxiety, report a significant relationship between anxiety level and perceptual rigidity under similar experimental conditions. These data lend strong support to the assumption that anxiety is associated with increased strength of prevailing perceptual expectancies.

If the preceding analysis is correct, the association between high levels of motivation and reduction in responsiveness to environmental cues could be attributed to the effects of motivation on the increased tendency toward perceptual rigidity under "stress" conditions. Therefore, one would expect a relatively high degree of manifest anxiety (i.e., drive) to be related to perceptual rigidity and, secondarily, to increased speed of closure. The purpose of this study was to explore this hypothesis. Specifically, in terms of the experimental tasks, it was predicted that relatively high anxiety subjects will: (a) tend to respond, on a decision-making task, prior to the availability of sufficient information to obtain the correct response (cognitive closure); (b) obtain relatively shorter latencies on a simple recognition task (perceptual closure); and (c) manifest more perceptual rigidity on a task where the initial stimulus of a series is successively modified to form a new stimulus configuration.

EXPERIMENTAL METHOD

Subjects

The total sample consisted of 127 fifth-grade children from a city school system.² The results are based on data from 118 subjects (54 female and 64 male) who completed the necessary tests on both testing occasions.

Anxiety Scale

The index of anxiety level for this study was derived from scores on the children's manifest anxiety scale (4). The reported test-retest reliability of the CMAS ranges from .61 to .75 depending on the age and sex of the Ss;

² Appreciation is extended to the officials of the Cedar Rapids, Iowa, school system, and the principals, teachers, and students of the Cleveland and Harrison schools, for their cooperation in this study. Special appreciation is also extended to Mr. Langdon Longstreth who participated in the collection of the data for this study.

an independent study (unpublished) by the present author yielded slightly higher test-retest reliabilities, over a three-months period, for seventh and eighth grade school children.

Perceptual-Cognitive Tasks

Decision Location Task. The index of speed of "cognitive closure" was derived from performance on the complex decision-making tasks (DLT-S) used in previous investigations concerned with the hypothesized relationship between "intolerance of ambiguity" and anxiety (14, 15). Each task consists of 15 cards that are presented in serial order to the Ss. The first card contains only a few elements of a picture which, by adding stimulus elements on successive cards, appears complete on card 15. A four-card sample task was administered for purposes of demonstration. Prior to the administration of each experimental series (tasks) Ss were presented with four "hypotheses" or possible correct responses. The hypotheses were drawn from the generic class of objects representing the correct response; e.g., the correct response on Task B was "cat," the pre-task hypotheses were, in order, dog, lion, cat, rabbit.

Perceptual Closure Task. The measure of speed of perceptual closure consisted of 12 pictures from the Street Gestalt Closure test (SGCT). A previous study (16) suggested that speed of closure tended to vary with the familiarity of the stimulus. Therefore, the pictures used here were categorized according to "difficulty level" on the basis of the data obtained in that investigation; four "easy," four "difficult," and four pictures of medium difficulty were included. The latency, to the nearest 1/10 second, of the first and of the correct response were recorded as the index of the speed of perceptual closure.

Perceptual Rigidity Task. An index of perceptual rigidity was obtained from a transition task (PRT) described in a previous study (16). The task was constructed by extending the series of 15 cards in each of the DLT-S tasks to a total of 30 stimulus cards. The cards from number 16 through 30 of transitional stages whereby the completed stimulus on card 15 (e.g., a cat) was modified until a new stimulus (e.g., a dog) appeared completed on card 30. Each transition task was administered immediately following its associated DLT-S task. As indicated below, S was instructed to report what he saw on each card *after* the *first response* had occurred in the DLT-S series. The measure of rigidity was the number of cards which elicited the response that was correct for card 15.

Procedure

The initial investigation compared the performance of 20 high and 20 low anxiety Ss on the three experimental tasks. A second experiment was conducted to explore the feasibility of group administration of the DLT-S and PRT tasks. The specific procedures for each experiment are presented below.

Experiment I. Ss scoring in the top and bottom 20 per cent of the distribution of scores on the CMAS were selected for the first experiment. Due to absences during the scheduled days of testing, the number of Ss in both groups was reduced to 20. Each S was first administered one of the DLT-S tasks (Task E; 14) and the associated transition task (Task E-T) from which the measures of "cognitive-closure" and perceptual rigidity were obtained. Prior to the administration of the task, a three-card sample series was shown to each S. Following the pre-task, S was permitted to look at the "hypothesis" card, on which the four possible correct responses for that task were printed, for 10 seconds. The instructions accompanying the DLT-S task emphasized that the purpose of the task was to determine *how few cards* the S needed to see in order "to put all the parts together" and report the correct answer.

Ss were permitted to view each card in the series for five seconds. If no response was made, E turned up the next card. The effect of pre-recognition responses on recognition was controlled as much as possible by requiring Ss to report what they saw on each card *after the first report* (*identification response*) occurred. E gave no indication to S that the first, or subsequent, responses (including the PRT) had been correct or incorrect. All responses, including those to the cards in the transition series (i.e., cards 16 to 30) were recorded verbatim. The trial of first response was recorded as the measure of the speed of perceptual-cognitive closure. The index of perceptual rigidity was the number of perseverative responses; i.e., responses made to the stimulus cards 16 to 30 that had been correct for card 15.

A modified form of the Street Gestalt test was administered following the DLT-S and PRT tasks. Each picture was presented to the Ss with instructions to identify the picture. If the first response was incorrect, Ss were asked to "try again." If the correct response was not elicited within two minutes, the next picture was presented to the subject. The latency of all responses was recorded to the nearest 1/10 of a second.

Experiment II. The total sample was administered a group form of three DLT-S and Transition tasks (Tasks A, B, C; 14) four months after Experiment I was completed. The stimuli were projected on a screen in the front of the room. The procedure and instructions were identical to those used in the individual sessions, with the exception that each stimulus card was exposed for only one second (controlled by leaf-diaphragm tachistoscope). Ss recorded their responses in booklets provided for that purpose. These booklets consisted of 15 pages; each page had a space for writing the response or indicating a "don't know" answer on each trial. If a response other than "dk" was made, Ss were instructed to indicate the certainty of their response on a three-point rating scale; i.e., uncertain, certain, very certain. After the exposure of each card and the recording of their response, Ss folded that page back and under; this, and the presence of teachers as proctors, precluded Ss from going back to change responses. Responses to

the stimuli 16 through 30 were recorded on the back of the booklets. The sample task was re-administered prior to the second and third tasks since the transition stimuli (PRT) were likely to introduce a different set than had prevailed initially on the first DLT-S task.

RESULTS

Experiment I

Hypotheses 1 and 2 stated that a high level of anxiety, under the present experimental conditions, would be associated with relatively stronger tendencies toward cognitive and perceptual closure as indicated by shorter "latency" on both the decision location task and the object recognition test. Table 1 presents the data relevant to these hypotheses.

TABLE 1
MEAN TRIAL OF FIRST DECISION AND LATENCY OF RESPONSE FOR
HIGH AND LOW ANXIETY SUBJECTS ON THE EXPERIMENTAL
MEASURES OF COGNITIVE AND PERCEPTUAL CLOSURE

	DLT-S Trial of First Response	SGCT DIFFICULTY LEVEL (latency in seconds)		
		Easy	Medium	Difficult
<i>High Anxiety:</i>				
Mean	3.45	1.64	9.51	24.88
Variance	11.85	.15	61.19	228.53
<i>Low Anxiety:</i>				
Mean	5.07	1.72	5.40	14.19
Variance	10.25	.45	10.71	150.43

The difference between high and low anxiety groups on the mean trial of first decision, on the DLT-S, (3.45 and 5.07 respectively) is statistically significant beyond the 10 per cent level of confidence ($t = 1.72$; $df = 38$; $.10 > p > .05$). Analysis of variance of the latency data obtained from performance on the SGCT task (Table 1) yielded an F ratio of 5.48 for the main effect of anxiety which, for 1 and 38 degrees of freedom, is significant between the .02 and .01 level of confidence. The main effect of difficulty level and the interaction between anxiety and difficulty were also statistically significant; ($F = 36.21$ and 21.32 ; $df = 2, 76$; $p < .001$). The high anxiety Ss obtained a relatively shorter latency of response only on the tasks classified as "easy"; over-all, this group had significantly longer latency of response. The lack of confirmation of the first hypothesis with respect to the more "difficult" perceptual tasks may be a function of

an artifact introduced by the procedures, or of the differential effects of drive on response strength in "simple" as compared to "complex" task situations (5, 7). This problem will be discussed more fully in a later section of the paper.

The third prediction was that individuals scoring higher on the CMAS would manifest relatively more perceptual rigidity as indicated by perseverative-response tendencies on the transition task (PRT). However, in order to attribute the response perseveration to *perceptual rigidity*, it is necessary to rule out the effect of reinforcement of verbal responses (i.e., number of correct responses) in the initial series of cards (DLT-S). The comparison of the high and low anxiety groups on the trial of correct response yielded a t of 1.63 ($df = 38$; $p > .10$). The product-moment correlations between number of correct responses prior to card 15 and perseverative response on cards 16 to 30 for the anxiety groups combined and separately also were statistically nonsignificant (High, $r = -.22$; Low, $r = -.27$; High + Low, $r = -.19$; $p > .10$). It would appear therefore, that any differences between the two groups on perseverative tendencies can be attributed to the effects of anxiety level on perceptual rigidity.

The mean number of perseverative responses was 6.90 for the high anxiety Ss and 5.01 for the low anxiety Ss. This difference is statistically reliable beyond the .05 level of confidence as indicated by a t of 2.31 ($df = 38$; $.05 > p > .02$). These data confirm the findings of a previous study concerned with the relationship between indices of anxiety level and response perseveration on this task (16).

Experiment II

The procedure for the second experiment did not permit administration of the Street Test (SGCT). However, two measures of cognitive closure were available; the trial of first response, and the certainty score associated with each pre-recognition response on the 15 card series of the DLT task. Furthermore, the group administration permitted the testing of sufficient number of Ss so that the performance of those scoring in the middle range of the CMAS could be examined. The following analysis was based on the performance of the Ss scoring in the upper, middle, and lower 20 per cent of the distribution of the CMAS score ($N = 24$ in each group). Table 2 presents the summary of the data obtained from each anxiety group on each of the three experimental DLT-S tasks.

Analysis of variance of these data indicated the differences between the anxiety groups, although in the predicted direction, did not reach a high level of reliability ($F = 2.03$; $df = 2, 69$; $.15 > p > .10$). A statistically significant difference was attributable to the task (trial) variable³ ($F = 4.36$; $df = 2, 138$; $.02 > p > .01$). Similar trends were obtained

³ The tasks were administered in the same order (A, B, C) to all groups.

TABLE 2

MEAN TRIAL OF FIRST RESPONSE ON THREE EXPERIMENTAL DLT-S
TASKS FOR HIGH, MEDIUM, AND LOW ANXIETY SUBJECTS

	Task A	Task B	Task C
<i>High Anxiety:</i>			
Mean	4.50	4.08	3.13
Variance	21.75	13.51	11.90
<i>Medium Anxiety:</i>			
Mean	6.08	5.88	4.79
Variance	15.89	11.53	12.67
<i>Low Anxiety:</i>			
Mean	6.21	5.92	4.63
Variance	15.32	14.04	12.00

in a previous study (14) and are probably a function of the repeated reinforcement of the "early" response set over a series of trials. Further analysis confirmed the impression gained from the table that most of the reliable differences reported above could be attributed to the difference between the high anxiety Ss as compared to the other two groups.

A second index of cognitive closure was obtained from the average degree of subjective certainty expressed by each S on the pre-recognition responses on each task. The average certainty score for each S was computed by dividing the sum of the weighted certainty scores (3 = Very Certain; 2 = Certain; 1 = Uncertain; 0 = Very Uncertain) by the number of trials between the *first decision* and the trial of *correct response*. Analysis of variance of these scores yielded an F ratio of 2.45 ($df = 2, 69$; $.10 > p > .05$) for the between-groups comparison. The main effect of tasks, and the interaction of anxiety level and tasks, did not prove to be statistically reliable.

Table 3 presents the data relevant to the hypothesized relationship between manifest anxiety and perceptual rigidity. Analysis of variance of these data indicated the differences between groups and tasks were both significant beyond the .02 level of confidence ($F = 5.44$, $df = 2, 69$; $F = 4.15$, $df = 2, 138$). The correlation between number of correct responses on the DLT-S and perseveration for the total group of Ss was $-.21$ which for 70 degrees of freedom is not statistically significant at the 10 per cent level of confidence. These findings are consistent with those obtained in the first experiment and lend support to the assumption that the perseveration effect is a function of perceptual, rather than response, rigidity.

CHILD DEVELOPMENT

TABLE 3

MEAN NUMBER OF PERSEVERATIVE RESPONSES ON THE THREE EXPERIMENTAL TASKS FOR HIGH, MEDIUM, AND LOW ANXIETY GROUPS

	Task A	Task B	Task C
<i>High Anxiety:</i>			
Mean	7.88	8.46	6.71
Variance	14.54	15.57	10.70
<i>Medium Anxiety:</i>			
Mean	7.00	5.62	4.12
Variance	12.42	10.32	5.94
<i>Low Anxiety:</i>			
Mean	6.54	6.00	5.11
Variance	8.42	13.83	7.75

DISCUSSION

These data indicate that high manifest anxiety, as measured by the children's anxiety scale, is associated with relatively more perceptual rigidity and, under certain conditions, increased speed of closure.⁴ The findings are generally consistent, therefore, with the experimental evidence on the effects of anxiety on perceptual-cognitive functioning (2, 10, 14, 15). Further, in view of the relationship between manifest anxiety and perceptual constriction or "stimulus fixation" (8, 10, 13, 18), the data indirectly support the assumption that "decreased responsiveness" to environmental cues (1, 2) may be some functional derivative of these perceptual consequences of high levels of motivation or drive.

Contrary to expectations, the high anxiety subjects obtained, over-all, a longer latency of response on the *perceptual closure* tasks than low anxiety subjects. These results, and the statistically significant interaction between anxiety level and task difficulty, might be attributed to one of two factors. First of all, the tasks were, unfortunately, administered in the order of difficulty. The relatively longer latency on the more difficult tasks may have been a function of the differential reaction to "failure" (incorrect responses) of the high-anxiety Ss. It had been noted in previous investigations (14, 16) that successive failure experiences, in general, tend to evoke a relatively stronger accuracy orientation on the part of anxious

⁴ There is no evidence from the results of this study to suggest these relationships could be attributed to differential levels of intelligence associated with high and low CMAS scores. The correlations between CTMM scores and CMAS, response perseveration, and latency of closure ranged from $-.17$ to $-.14$; none was significantly different from zero.

subjects. Response suppression may, therefore, account for the negative findings with respect to this task.

One might also argue that the obtained relationship between manifest anxiety and performance on these tasks is indicative of the effects of anxiety on strength of the response system rather than on the mediating functions of the perception (6). According to the general laws of associative learning, performance on a "perceptual" task may be considered a joint and multiplicative function of the strength of the perceptual response disposition (associative strength) and drive level (5, 7). The obtained relationship between manifest anxiety and response perseveration on the transition task (PRT) would, therefore, be a consequence of the tendency for the dominant response tendency (correct response) to be evoked by more dissimilar stimuli. Further, the interaction between anxiety level and task difficulty on the Street closure test also would be predicted if one made the additional assumption that "easy" and "correct response dominant" are synonymous.

However, a recent experiment reported by Saltz and Hoehn (12) failed to lend support to the latter assumption. Further, the data reported here indicated that the trial of correct response on the DLT-S task did not differentiate between high and low anxiety subjects, nor was the frequency of correct response significantly related to the degree of perseveration on the transition task. As suggested in the earlier discussion, these data lend support to the contention that perceptual rigidity, associated with high anxiety levels, is a more important determinant of perseveration tendencies than response (associative) strength.

While the results of this study permit only the most cautious generalization, when considered in conjunction with other recent investigations (2, 3, 10, 16), the data appear most consistent with the assumption that motivational variables exert certain direct influences on perceptual processes. Despite the controversy surrounding this assumption, both theoretical and empirical considerations indicate the distinction between the *perceptual* system and other response systems (e.g., 2, 6, 11) is a valid one. Further research, directed toward determining the conditions under which motivational factors influence perception, should permit more specific generalizations concerning the nature of these effects. Further, in view of the recent evidence concerning the role of early perceptual experience on subsequent behavior tendencies, an analysis of the extent to which the perceptual consequences of motivation (e.g., anxiety) varies with maturational and experiential factors would seem particularly important for achieving a more adequate understanding of psychological development.

SUMMARY

The purpose of this study was to determine some of the perceptual consequences of manifest anxiety in children. It was proposed that the reduction in behavioral adaptivity associated with high levels of motivation

might be due to the effects of certain motivational states on perceptual rigidity and closure tendencies. Therefore, individuals with relatively high levels of manifest anxiety (drive) would be expected to achieve a shorter latency on a cognitive and perceptual closure task and to respond to a previously reinforced stimulus (perception) more frequently than individuals with a lower level of anxiety.

The children's form (CMAS) of the Taylor manifest anxiety scale was administered to 118 fifth-grade children. Measures of cognitive-perceptual closure and rigidity were individually administered to the upper and lower 20 per cent of those Ss (Experiment I). Four months later a second series of tasks were administered to the total group of Ss and the results obtained from the upper, middle, and lower 20 per cent of the CMAS were analyzed (Experiment II). The results generally confirmed the predicted relationships between manifest anxiety and perceptual rigidity and closure phenomenon.

REFERENCES

1. AXELROD, H. S., COWEN, E. L., & HEILIZER, F. Some correlates of manifest anxiety in stylus maze learning. *J. exp. Psychol.*, 1956, 51, 131-138.
2. BRUNER, J. A. On perceptual readiness. *Psychol. Rev.*, 1957, 64, 123-152.
3. BRUNER, J. A., MATTER, J., & PAPNEK, M. L. Breadth of learning as a function of drive level and mechanization. *Psychol. Rev.*, 1955, 62, 1-10.
4. CASTANEDA, A., PALERMO, D., & McCANDLESS, B. R. The children's form of the manifest anxiety scale. *Child Developm.*, 1956, 27, 317-326.
5. CASTANEDA, A., PALERMO, D., & McCANDLESS, B. R. Complex learning and performance as a function of anxiety in children and task difficulty. *Child Developm.*, 1956, 27, 327-332.
6. ERIKSEN, C. W., & BROWN, C. A. An experimental and theoretical analysis of perceptual defense. *J. abnorm. soc. Psychol.*, 1956, 52, 224-230.
7. FARBER, I. E., & SPENCE, K. W. Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.*, 1953, 45, 120-125.
8. GILCHRIST, J. C., LUDEMAN, J. F., & LYNCH, W. Values as determinants of word recognition thresholds. *J. abnorm. soc. Psychol.*, 1954, 49, 423-426.
9. HEILIZER, F., AXELROD, H. S., & COWEN, E. L. The correlates of manifest anxiety in paired associate learning. *J. Pers.*, 1956, 24, 463-474.

CHARLES D. SMOCK

10. MOFFIT, J. W., & STAGNER, R. Perceptual rigidity and closure as functions of anxiety. *J. abnorm. soc. Psychol.*, 1956, 52, 354-357.
11. NEISSER, U. Experimental distinction between perceptual process and verbal response. *J. exp. Psychol.*, 1954, 47, 399-402.
12. SALTZ, E., & HOEHN, A. J. A test of the Taylor-Spence theory of anxiety. *J. abnorm. soc. Psychol.*, 1957, 54, 114-117.
13. SILVERMAN, R. E., & BLITZ, B. Learning and two kinds of anxiety. *J. abnorm. soc. Psychol.*, 1956, 52, 301-303.
14. SMOCK, C. D. The influence of psychological stress on the perception of incongruity. *J. abnorm. soc. Psychol.*, 1955, 50, 354-356.
15. SMOCK, C. D. The influence of psychological stress on the intolerance of ambiguity. *J. abnorm. soc. Psychol.*, 1956, 50, 177-182.
16. SMOCK, C. D. The relationship between intolerance of ambiguity, generalization and speed of perceptual closure. *Child Developm.*, 1956, 28, 27-36.
17. SMOCK, C. D. The relationship between test anxiety, threat-expectancy and recognition thresholds for words. *J. Pers.*, 1956, 25, 191-201.
18. SMOCK, C. D. Recall of interrupted and non-interrupted tasks as a function of experimentally induced anxiety and motivational relevance of the task stimuli. *J. Pers.*, 1957, 25, 589-599.
19. TAYLOR, JANET A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, 48, 285-290.

THE CAUSATION AND MANIFESTATIONS OF EMOTIONAL BEHAVIOR IN FILIPINO CHILDREN¹

RACHEL MARZAN INSELBERG²
Columbus, Ohio

Many studies on the emotional behavior of children have dealt with aggressive behavior. A small number of investigators have attempted to survey the wide spectrum of emotional reactions of children.

The objectives of the present study were to compare two age groups of Filipino children with respect to (a) the situations causing emotional episodes, (b) the manifestations of emotional behavior, and (c) the reactions of adults to these emotional episodes.

Blatz, Chant, and Salter (2) report that as the child grows older emotional behavior undergoes a change, becoming increasingly directed and correspondingly less irrelevant and chaotic in nature.

Several investigators have studied aggressive behavior in children (1, 3, 4, 5, 9). Doob and Sears (4) found that overtness of aggression varied inversely with the amount of punishment anticipated. The findings of Walters, Pearce, and Dahms (9) indicated that aggression tended to increase from two through four years of age. Appel (1) found that a major number of aggressive episodes arose from a desire for the possession of property. This interest in things per se tended to decrease with age.

Dollard *et al.* (3) concluded that the peak of negativism occurring at about three years of age resulting from frustration caused by feeding problems, cleanliness, and training. Frederiksen (5) reported that frustration brought about negativistic behavior of the withdrawal type.

Heathers (7) and Hattwick and Sanders (6) found that as a preschool child grows older, he increasingly seeks attention and approval from adults.

¹ This study was conducted at the Philippine Women's University. The author wishes to acknowledge the inspiration derived from Mrs. Doreen Gamboa, Head, Department of Child Development.

² Formerly connected with Philippine Women's University.

CHILD DEVELOPMENT

SUBJECTS

The subjects studied were 62 Filipino children, 31 from two to four years of age and an equal number between five and seven years of age. The majority of the children belonged to the families of the observers. Thirty-six per cent of the subjects belonged to families of the professional group (e.g., lawyers, engineers, physicians, and teachers); 36 per cent, to the laboring group; and 28 per cent, to government employees, secretaries, bookkeepers, and owners of small retail stores.

PROCEDURE

The method used in this study was direct observation with predetermined categories. The classification scheme used was based partly on the work of Landreth (8). Repeated revision yielded the final instrument employed in the study. A check list was used to facilitate observation.

Emotional episodes were studied from three standpoints: situations causing emotional behavior, manifestations of emotional behavior, and reactions of adults in dealing with these episodes.

This study was mainly concerned with the direct or immediate situations causing specific outbursts. Hence, such contributing factors as ill-health, fatigue, and time of day were not considered.

Observations were carried out by teachers attending summer school. The observations were made at home over a six-week period. The observers were oriented as to the objectives of the study, and were asked to record pertinent information following each emotional episode occurring in the home.

A grand total of 755 emotional episodes were recorded during the observation period, 390 in the two to four year group and 365 in the five to seven year group. It should be noted that the number of manifestations of emotional behavior and of adult reactions exceeded the number of emotional episodes, since more than one manifestation or adult reaction may have been associated with a particular emotional episode.

RESULTS

The frequencies of emotional episodes caused by various situations are presented in Table 1. The corresponding frequencies of the two age groups are compared by means of a *t* test. The level of significance is given in the table.

Established home and health routine gave rise to a significantly larger frequency of emotional episodes in the older group. On the other hand, play activities resulted in a significantly larger frequency of emotional episodes in the younger group. There was no significant difference in the

TABLE I

SITUATIONS CAUSING EMOTIONAL BEHAVIOR—FREQUENCIES OF EMOTIONAL EPISODES

	2-4 YEAR GROUP			5-7 YEAR GROUP			
	No.	Freq. (%)	Rank	No.	Freq. (%)	Rank	Signif. Level
<i>Established Home and Health Routine:</i>							
Eating	61	15.6		63	17.3		ns
Sleeping	18	4.6		17	4.7		ns
Working (errands, chores)	13	3.3		24	6.6		0.05
Grooming	12	3.1		33	9.0		0.001
Bathing	8	2.0		12	3.3		ns
TOTAL ..	112	28.7	2	149	40.8	1	0.001
<i>Play Activities:</i>							
Conflict with children	91	23.3		69	18.9		ns
Conflict with adults	62	15.9		32	8.8		0.05
Frustration with toys & materials	15	3.8		5	1.3		0.05
Frustration by animals	2	0.5		1	0.3		ns
TOTAL ..	170	43.6	1	107	29.3	2	0.001
<i>Other Situations:</i>							
Separation from adults	53	13.6		25	6.9		0.005
Deprivation of various objects	19	4.9		26	7.1		ns
Favoritism	9	2.3		2	0.5		0.05
Medical & dental care	8	2.0		14	3.8		ns
Fear—strangers & unfamiliar objects ..	8	2.0		2	0.5		ns
School	3	0.8		23	6.3		0.001
Teasing	6	1.5		1	0.3		ns
Church	2	0.5		6	1.6		ns
Performing for visitors	0	0		10	2.7		0.005
TOTAL ..	108	27.7	2	109	29.9	2	ns
GRAND TOTAL ..	390	100.0		365	100.0		

total frequency of emotional episodes caused by the miscellaneous situations grouped under "Other Situations." Inspection of these data indicates that some individual situations listed under one of the three broader categories gave rise to a significantly higher frequency of episodes in one of the two age groups. Several situations, on the other hand, resulted in essentially the same frequency of episodes in the two age groups.

A binomial test provided the basis for ranking the three general groups of situations according to the frequencies of emotional episodes that they gave rise to within each age group. Groups of situations receiving the same rank did not differ significantly in the frequencies of emotional episodes ($p < .05$).

The frequencies of the various manifestations of emotional behavior in the two age groups of children are presented in Table 2. A t test was used in comparing the corresponding frequencies of the two age groups. A

CHILD DEVELOPMENT

TABLE 2

MANIFESTATIONS OF EMOTIONAL BEHAVIOR IN TWO AGE GROUPS OF FILIPINO CHILDREN

	2-4 YEAR GROUP			5-7 YEAR GROUP			Signif. Level
	No.	%	Rank	No.	%	Rank	
Vocal (sobbing, loud cries, screaming) ..	358	52.4	1	117	24.7	2	0.001
Overt physical aggression against others (kicking, boxing, beating, pinching) ..	137	20.0	2	139	29.3	1	0.001
Overt physical aggression against self (pulling hair, tearing clothes, banging head)	38	5.6	3	20	4.2	3	ns
Withdrawal (reporting, murmuring, running away, trembling, grimacing, looking ashamed)	103	15.1	2	176	37.2	1	0.001
Relatively passive positive reactions (following, clinging, insistence, asking for help)	47	6.9	3	22	4.6	3	ns
TOTAL ..	683	100.0		474	100.0		

binomial test was used in ranking the five general groups of manifestations as to relative importance within each age group.

The frequency of vocal manifestations was much greater in the two to four year group. The emotional behavior of the five to seven year group, on the other hand, was characterized by significantly greater frequencies of overt physical aggression against others and withdrawal. It should be noted that the frequency of behavior of the withdrawal type was practically two and a half times as great in the older group as that encountered in the younger group. There were no significant differences in the frequencies of

TABLE 3

REACTIONS OF ADULTS

	2-4 YEAR GROUP			5-7 YEAR GROUP			Signif. Level
	No.	%	Rank	No.	%	Rank	
Indulgence (giving in, bribing, compromising, promising, flattering, enticing, etc.)	276	58.9	1	220	45.2	1	0.001
Verbal (threatening, scolding, commending, shouting)	97	20.7	2	137	28.1	2	0.01
Overt physical aggression toward child (slapping, pinching, forcing, dragging)	60	12.8	3	95	19.5	3	0.01
Other disciplinary measures (deprivation, isolation, favoritism)	11	2.3	4	23	4.7	4	0.05
Relatively passive reactions (ignoring, staring, noninterference)	25	5.3	4	13	2.7	4	0.05
TOTAL ..	469	100.0		488	100.0		

overt physical aggression against oneself and relatively passive positive reactions.

The reactions of adults to the emotional behavior of children are presented in Table 3. In this study, the adults included parents, relatives, and domestic help that were in close contact with the children at the time that emotional episodes took place. A *t* test was used in making the comparisons between the two age groups, while a binomial test was employed in ranking the groups of adult reactions within each age group.

The adults were significantly more indulgent toward the younger group. Conversely, they tended to be significantly more verbal and exercised more overt physical aggression and disciplinary measures toward the older group.

DISCUSSION

The results of the present study can be better understood when viewed against the background of Filipino culture. The prevailing authoritarian pattern places great emphasis on obedience. Respect for elders is often carried to the extreme of prohibiting expression of opinions by children. Consequently, submissiveness is equated with good behavior.

It should not seem contradictory, however, that in spite of the above considerations allowances are made for age. Adults tend to be more indulgent toward younger children, as is brought out in Table 3. As children grow older and approach school age, they are increasingly expected to conform to the dictates of their elders. Older children are treated as miniature adults; they are expected to conform to a pattern of being seen but not heard.

As the child grows older, training in established home and health routine is emphasized. This is illustrated by the increasing frequency with age of emotional episodes attributable to established home and health routine (Table 1).

The increased consciousness with age of intense adult disapproval of quarreling helps explain the findings presented in Table 1, in connection with play activities. Those data indicate that emotional episodes arising from conflicts with children and with adults tend to decrease with age, the difference being statistically significant only in the case of conflicts between children and adults.

It was brought out (Table 2) that as the child grew older the manifestations of emotional behavior became increasingly directed and correspondingly less irrelevant in nature. The older child was significantly less vocal and considerably more aggressive than the younger one. Superimposed on this series of changes, the pressures exerted on older children as a result of cultural values and norms are much more pronounced. Fear of punishment for being aggressive, as well as an effort to win adult approval, apparently account for the increased frequency of substitute behavior, such as withdrawal behavior encountered in the older group.

CHILD DEVELOPMENT

SUMMARY

The objectives of the present study were to compare two age groups of Filipino children, two to four and five to seven years of age, with respect to (a) the situations causing emotional episodes, (b) the manifestations of emotional behavior, and (c) the reactions of adults to these episodes.

The major findings of the study were:

1. Established home and health routine gave rise to a significantly larger frequency of emotional episodes in the five to seven year group.

2. Play activities resulted in a significantly larger frequency of emotional episodes in the two to four year group.

3. The emotional behavior of the two to four year group was primarily manifested vocally while in the five to seven year group emotional behavior was characterized by withdrawal and overt physical aggression against others.

4. The adults were significantly more indulgent toward the younger group.

The results are interpreted against the background of Filipino culture.

REFERENCES

1. APPEL, M. H. Aggressive behavior of nursery school children and adult procedures in dealing with such behavior. *J. exp. Educ.*, 1943, 11, 185-199.
2. BLATZ, W. E., CHANT, S. N. F., & SALTER, M. D. Emotional episodes in the child of school age. *Univer. Toronto Studies, Child Develpm. Series*, 1937, 9.
3. DOLLARD, J., DOOB, L. W., MILLER, N. E., MOWRER, O. H., & SEARS, R. R. *Frustration and aggression*. New Haven: Yale Univer. Press, 1939.
4. DOOB, L. W., & SEARS, R. R. Factors determining substitute behavior and the overt expression of aggression. *J. abnorm. soc. Psychol.*, 1939, 34, 293-313.
5. FREDERIKSEN, N. The effects of frustration on negativistic behavior of young children. *J. genet. Psychol.*, 1942, 61, 203-226.
6. HATTWICK, L. A., & SANDERS, M. K. Age difference in behavior at the nursery school level. *Child Develpm.*, 1938, 9, 27-47.
7. HEATHERS, G. Emotional dependence and independence in nursery school play. *J. genet. Psychol.*, 1955, 87, 37-57.
8. LANDRETH, C. Factors associated with crying in young children in the nursery school and the home. *Child Develpm.*, 1941, 12, 81-97.
9. WALTERS, J., PEARCE, D., & DAHMS, L. Affectional and aggressive behavior of preschool children. *Child Develpm.*, 1957, 28, 15-26.

PHYSIOLOGICAL INSTABILITY DURING ADOLESCENCE

DOROTHY H. EICHORN *and* JOHN P. McKEE

University of California, Berkeley

Several investigators (6, 18, 27, 29) have suggested that adolescence is characterized by a temporarily increased physiological instability. Bierring's (5) cross-sectional data for preadolescent and adolescent boys imply the contrary, for they show decreasing coefficients of variation as age increases. This paper considers the question of physiological instability during adolescence by examining longitudinal data from the Adolescent Growth Study (19). In addition, it summarizes previously published findings about physiological instability at other ages.

PREVIOUS FINDINGS

Basal Metabolic Rate

At the time normative standards for basal metabolism were established, a few investigators attempted to evaluate individual variation by making repeated measurements on the same individuals. The number of subjects in most of these studies was small. Berkson and Boothby (4) utilized 15 consecutive measures of basal metabolism made over periods of not longer than three months on 10 adult females and 23 adult males. The mean of the distribution of the individual standard deviations was 1.33 cal./sq.m./hr. for males and 1.61 for females; the corresponding coefficients of variation were 3.5 per cent and 4.7 per cent. In the same paper these writers give the standard deviations and coefficients of variation for cross-sectional measures of many hundreds of adults as 2.58 cal./sq.m./hr. for males (C.V. = 6.7 per cent) and 2.42 cal./sq.m./hr. for females (C.V. = 6.9 per cent). Berkson and Boothby also recalculated data from a number of earlier studies which included repeated measures on from one to 17 subjects. In general, the individual variabilities so calculated were of the

same order as the ones for their own subjects. Benedict (3) reported the individual ranges in BMR of 35 adult males tested on from five to 53 different days over periods of from five days to four years and five months. Expressed as proportion of the minimum rate, the individual ranges varied from 3.5 per cent to 31.3 per cent, with a mean of 13.9 per cent. Harris and Benedict (15) calculated the coefficients of variation for distributions of BMRs on each of 11 adult males tested on from 20 to 53 different days. The individual coefficients of variation ranged from 2.3 per cent to 5.3 per cent, with an average of 3.9 per cent. For periods not exceeding 10 to 15 days, the magnitude of the variation was positively correlated with the interval between the observations. Freeman and Giffin (8) made three determinations of the BMR of 50 adult males on two different occasions. The two occasions were often widely separated in time. The test-immediate retest correlation was .76; the test-later retest correlation was .29. Griffith (10) found the average variation between test and immediate retest of BMR to be 1.2 per cent for five adult subjects; the average daily variation from the monthly means was 2.5 per cent.

Herrington (16) took daily measures of several physiological functions on 11 young adult males for 45 days. For BMR, the individual standard deviations ranged from 3.2 to 5.5 cal./hr. The correlation between the individual means and standard deviations was $.10 \pm .30$, indicating no relationship between the level and the variability.

Harmon (13) attacked the problem of inter- and intra-individual variability in metabolic rate with an analysis of variance of 12 measures on each of 29 young adult males. Tests were made at 8:00 and 12:00 A.M., and 6:00 and 10 P.M. on each of three different days about a week apart. For basal measures (8:00 A.M.) the between subjects variance was 19.6 cal./sq.m./hr. and the within subject variance (for measures taken at the same time of day on different days) was 4.4 cal./sq.m./hr. Contrary to expectation, the basal determinations were not significantly less variable than other determinations. The test-retest correlations for the three days averaged .56. Harmon concluded that measures of basal metabolic rate are not highly reliable even under the best conditions. One draws similar, though less extreme, conclusions from the paper by Peeler *et al.* (25) which reports correlations of about .65 for measures separated by one to six weeks. In general most, but not all, previous reports (12, 21, 32) of individual variability in BMR have reported an average somewhere between 4 and 8 per cent.

Body Temperature

In another article, using the same young adult subjects, Harmon (14) applies the same analysis of variance approach to the question of variability in oral temperature. The standard deviations for comparable readings upon different individuals at the same time of day and for readings made upon the same individuals at the same hour on different days were about $.6^{\circ}$ F.

Kleitman and Ramsaroop (22) studied oral temperature variations in six adult males, four adult females, and two adolescent females. The subjects took their own temperatures every two to four hours over considerable periods of time (e.g., 182 days). For a given individual the range for any one hour was often greater than the diurnal variation. The basal temperatures were the least variable, both within and between subjects; individual ranges at 6:00 A.M. varied from 1.3 to 1.5° F. Wright (33) presents graphs of one individual's oral temperatures at 4:00, 8:00, and 12:00 A.M. and 4:00, 8:00, and 12:00 P.M. over a period of three days. The range of temperature at any one hour varied from about .4 to .8° F.

Bayley and Stolz's (2) data for early childhood show increasing stability in body temperature up to age three. The most closely comparable data for ages two to four (23) show no systematic age trend in variability. Comparable data for later childhood and adolescence are not available, but for adults, Horvath *et al.* (17) report the mean individual range for rectal temperatures as 1.2° F. for males and 1.5° F. for females. Their data, obtained under approximately basal conditions, are based on 29 consecutive daily readings on 16 adult males and 38 adult females. The implication is that a fair degree of variability remains present even in adulthood.

Blood Pressure

The data on individual variability in blood pressure are even more scanty than those on BMR and temperature. Grollman (11) obtained blood pressure readings on one subject of 98/71 on one day and 102/70 four days later. After eight months the same subject's pressures were 100/70 on one day and 98/73 three days later. In a paper dealing with the relationship of auscultatory and intra-arterial pressures, Ragan and Bordley (26) report auscultatory measures taken over several days for each of four adult males. The individual ranges for systolic pressure varied from four to 10 mm. Hg.; for diastolic pressure, from two to six mm. Hg. In Herrington's study the standard deviations of the individual distributions of the systolic blood pressures of 11 young adult males ranged from 3.57 mm. to 6.93 mm. over a period of 45 days. The correlation of the means and standard deviations was .13, indicating no relationship between level of systolic blood pressure and variability. Graham, Hines and Gage (9), who followed several hundred children over a 10-year period, report an increasing standard deviation with increasing chronological age. Unfortunately they analyzed their data cross-sectionally rather than longitudinally so that intra-individual variability cannot be determined.

Pulse Rate

The individual standard deviations for the pulse rates of Herrington's subjects ranged from 2.6 beats/min. to 5.5 beats/min. The correlation between means and standard deviations was .42. Herrington comments,

"It is not surprising to find that the pulse rate, which is certainly the most labile of these functions, shows a somewhat higher relation than the remaining variables" (16, p. 134).

Grollman (11) reports repeated pulse rates of two subjects, one measured under stringent laboratory conditions, the other under less controlled conditions. For the first subject, differences between measures separated by three days ranged from 0 to 4 beats; over a period of a year and a half the range was 8 beats/min. For the second subject, the range of differences between measures separated by three days was 10 beats/min.; over a period of 10 months the range was 12 beats/min. Palmer *et al.* (24) counted the pulse rate of one adult male on six consecutive days. The range for these six days was 4 beats/min.; the differences between any two successive days ranged from .5 to 2.0 beats/min.

Sutliff and Holt (31), in a summary of early data that is still used as a standard for textbooks (18), remark that "variability of pulse rate under basal conditions is greater in children than adults, and greater in females than males." Since the early material was almost exclusively cross-sectional, the question of age differences within individuals remains unsettled.

Summary of Previous Findings

The single most appropriate summarizing statement is that for all the measures discussed the amount of information about individual variation is woefully small—so small as to give little ground for generalizing about age trends. Despite this fact, and despite the fact that data have been analyzed in such a variety of ways that comparisons are very difficult, a few generalizations about BMR and body temperature are possible.

For BMR the correspondence between pairs of measures decreases as the interval between measures increases (8, 10, 13). For intervals up to about three months the coefficient of variation for BMR measures of the same individuals appears to lie somewhere between 3 and 5 per cent (4, 15). This is in the neighborhood of 1 to 2 cal./sq.m./hr. Griffith's five cases and Freeman and Giffin's correlation of .76 suggest that the immediate test-retest reliability of BMR is satisfactory. But Harmon's correlation of .41 for a two-week interval and Freeman and Giffin's .29 for a longer period indicate that even in adults many changes are to be expected in an individual's BMR when the measures are separated by any considerable period of time.

For body temperature Bayley and Stolz's data show increasing stability during the first three years, but the individual ranges in basal body temperature for Kleitman and Ramsaroop's older subjects varied from 1.3° to 1.5° F. And Horvath *et al.* report *mean* ranges of about the same magnitude. The individual variability of measures taken under other than basal conditions is even larger. Thus, just as for BMR, we may expect even adults with their hypothetical "stability" to show a rather sizable variation in body temperature.

For pulse rate and blood pressure the only generalization justified by the data so far published is that some individuals vary more than others!

THE PRESENT INVESTIGATION

Subjects and Procedure

Fifty boys and 50 girls from the physiological panel of the Adolescent Growth Study¹ (19) constitute the sample. Insofar as possible, physiological measures of each subject were taken at six-month intervals between the ages of 11.5 and 17.5 years. Of the many physiological indices available, those which seemed most likely to exhibit intra-individual changes in stability associated with adolescence were BMR, body temperature, systolic and diastolic blood pressure, and pulse rate. At the semiannual testings, BMR was determined three times on each of two successive days; basal oral temperature was taken at least once on two successive days; blood pressures and pulse rates were taken after each of the six metabolic determinations.²

Methods of Analysis

Evidence for systematic change in intra-individual variability was sought in two ways: first, by determining the range of each subject's measures for both testing days combined and plotting the mean range against age, and second, by determining for each subject the absolute difference between the medians of his measures on the two successive testing days and plotting the mean absolute difference against age. The first procedure should reveal any general change in physiological stability during the adolescent period; the second is aimed at determining whether such a change is primarily a change in diurnal variation. While it is unlikely, there could be systematic change in diurnal variation without a more general effect.

Results

Mean range. Figures 1 and 2 give, for each sex separately, the mean range by age for BMR, oral temperature, pulse rate, and systolic and diastolic blood pressure.³ Except for temperature, the curves for the two sexes are rather similar for any given physiological function, but the age trends are not the same for all functions. For BMR intra-individual variability declines irregularly, while for pulse rate and the blood pressures there is an increase. For temperature, the boys' variability decreases errat-

¹ Harold E. Jones directed the investigations. Nathan W. Shock supervised the physiological panel.

² Details of the procedures are given in (29).

³ Because of missed examinations the *N* for most of the points on the various graphs is typically nearer to 40 than to 50. At the extreme age groups (11.5 and 17.5 years) the *N*s are small (13 to 24). Points with *N*s of less than 20 have been indicated by dotted lines on the graphs.

CHILD DEVELOPMENT

ically and the girls' curve is so unsystematic that chance appears to be the most reasonable explanation. In general, the curves fall into two groups. Those for pulse and the blood pressures tend to parallel each other as do the curves for BMR and oral temperature. The latter parallel is somewhat more marked for boys than for girls. This is not surprising since we have already observed (7) that, for these subjects at least, body temperature

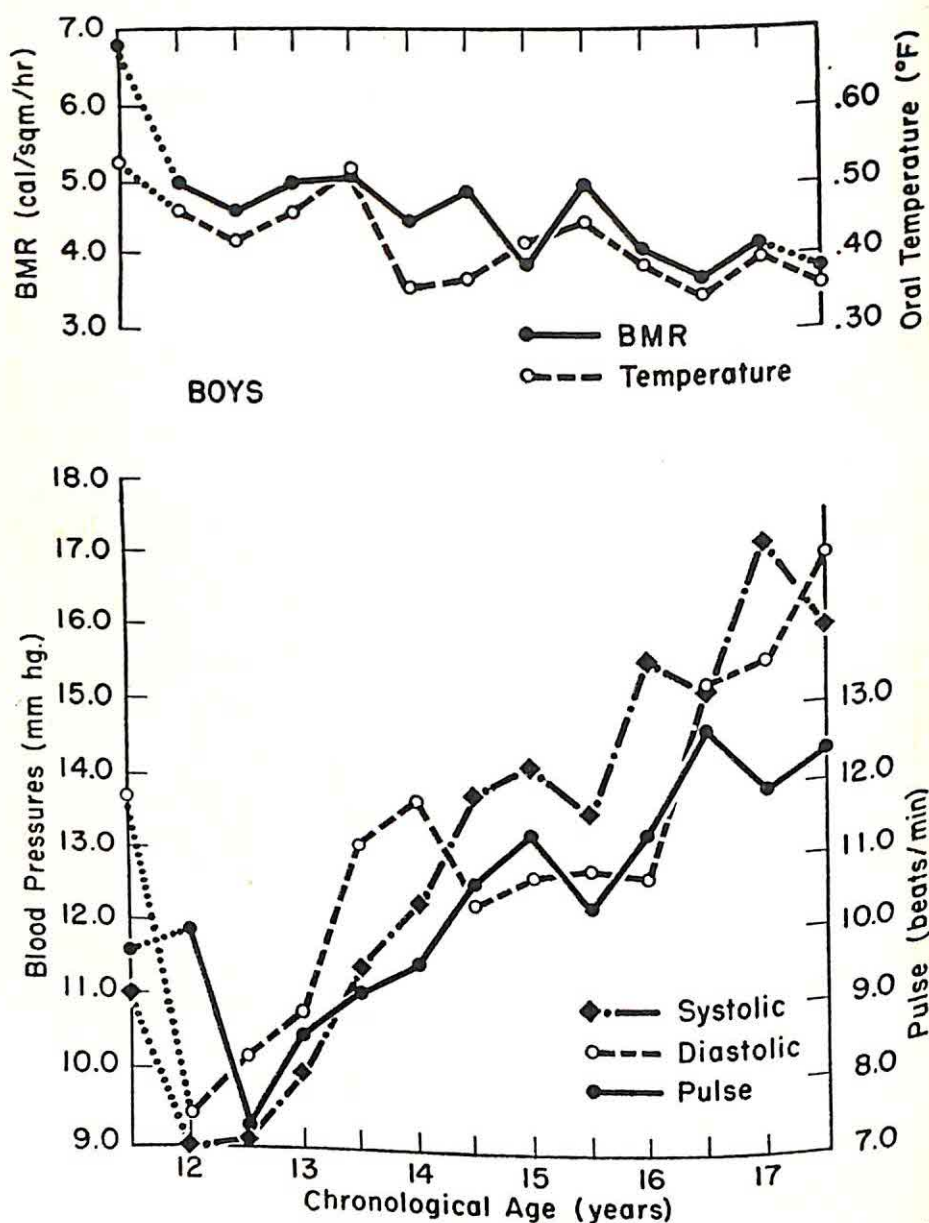
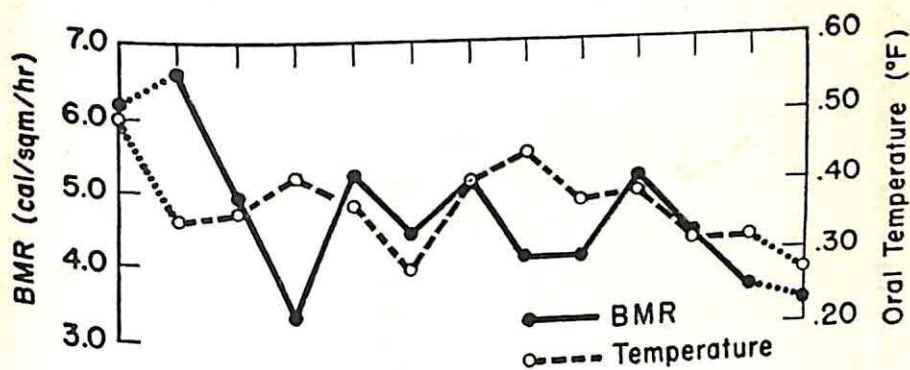


FIGURE 1—Mean range for basal metabolic rate, oral temperature, pulse rate, and systolic and diastolic blood pressure as a function of age—boys.

and BMR are correlated in boys and uncorrelated in girls. Although more of the curves in the two figures show a rise than show a fall, the data do not support the notion that there is a *generalized* physiological instability during adolescence. Rather they suggest that some functions become more stable and others less. Nor do the data support the view that such a hypothetical instability is *temporary*, for in no case does the curve first rise and then fall. It may be, however, that a fall would appear if the data included age 20 and beyond. While the number of adult cases for



GIRLS

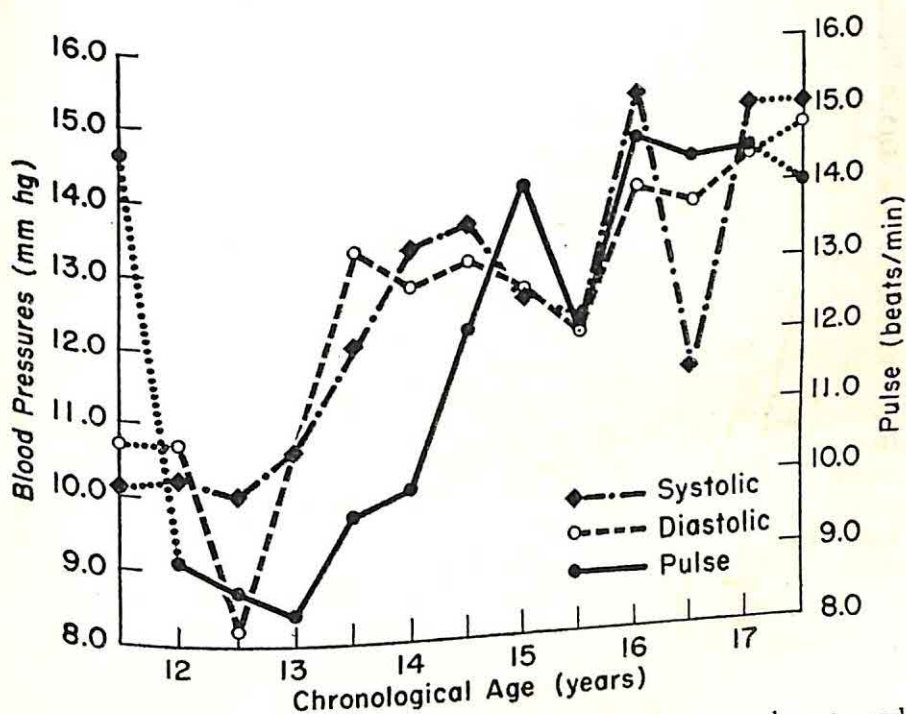


FIGURE 2—Mean range for basal metabolic rate, oral temperature, pulse rate, and systolic and diastolic blood pressure as a function of age—girls.

CHILD DEVELOPMENT

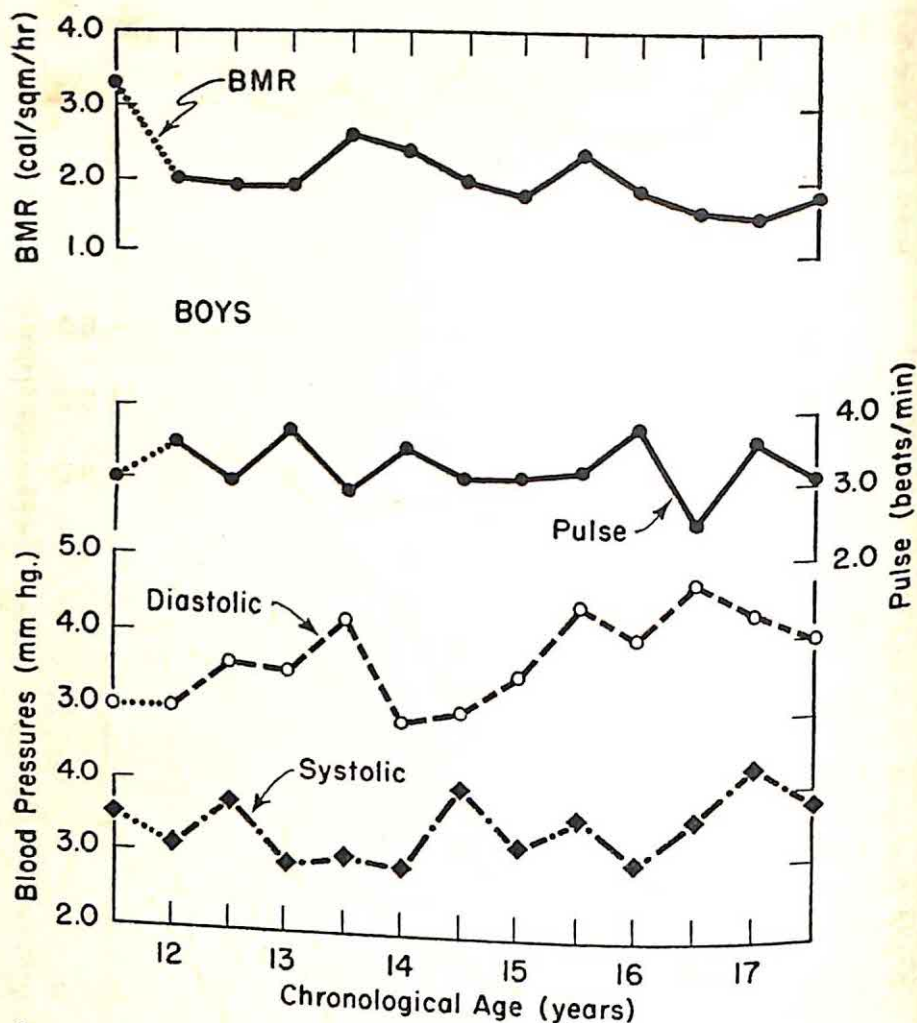


FIGURE 3—Mean absolute difference between medians on Day 1 and Day 2 for basal metabolic rate, pulse rate, and systolic and diastolic blood pressure as a function of age—boys.

whom day to day ranges are available is extremely small, the values reported (11, 24, 26) are somewhat smaller than the values we have obtained for pulse rate and blood pressures.

Diurnal variation. Figures 3 and 4 give, for each sex separately, the age trend in the mean absolute difference between the measures taken on day 1 and those taken on day 2.⁴ Of the eight curves in these two figures only three show any evidence of systematic change. Among girls there is

⁴ There is no figure for body temperature. For many cases only one measure of temperature was taken on each of the two days. To have omitted them would have reduced the number of cases too severely. Consequently we have used only the first measure taken on each day which makes the range identical with the difference between days.

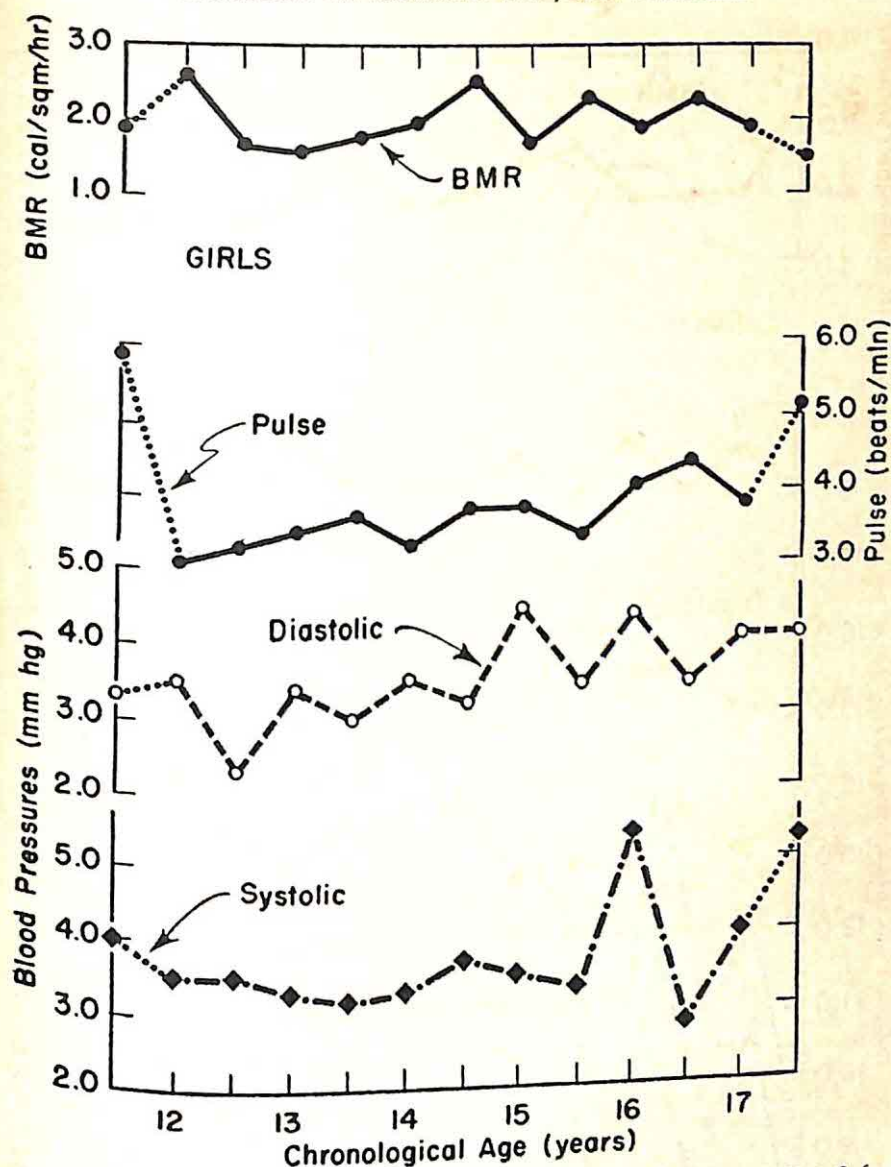


FIGURE 4—Mean absolute difference between medians on Day 1 and Day 2 for basal metabolic rate, pulse rate, and systolic and diastolic blood pressure as a function of age—girls.

a definite but small increase in the diurnal variation in pulse rate. There also appears to be a slight upward trend in the variation in diastolic pressure for girls and a slight decrease in BMR for boys. Otherwise, the curves are so irregular that they suggest no more than chance fluctuations from a horizontal slope. As was true of the curves for the range, these curves of diurnal variation suggest that the answers to questions about physiological stability during adolescence depend upon the particular func-

CHILD DEVELOPMENT

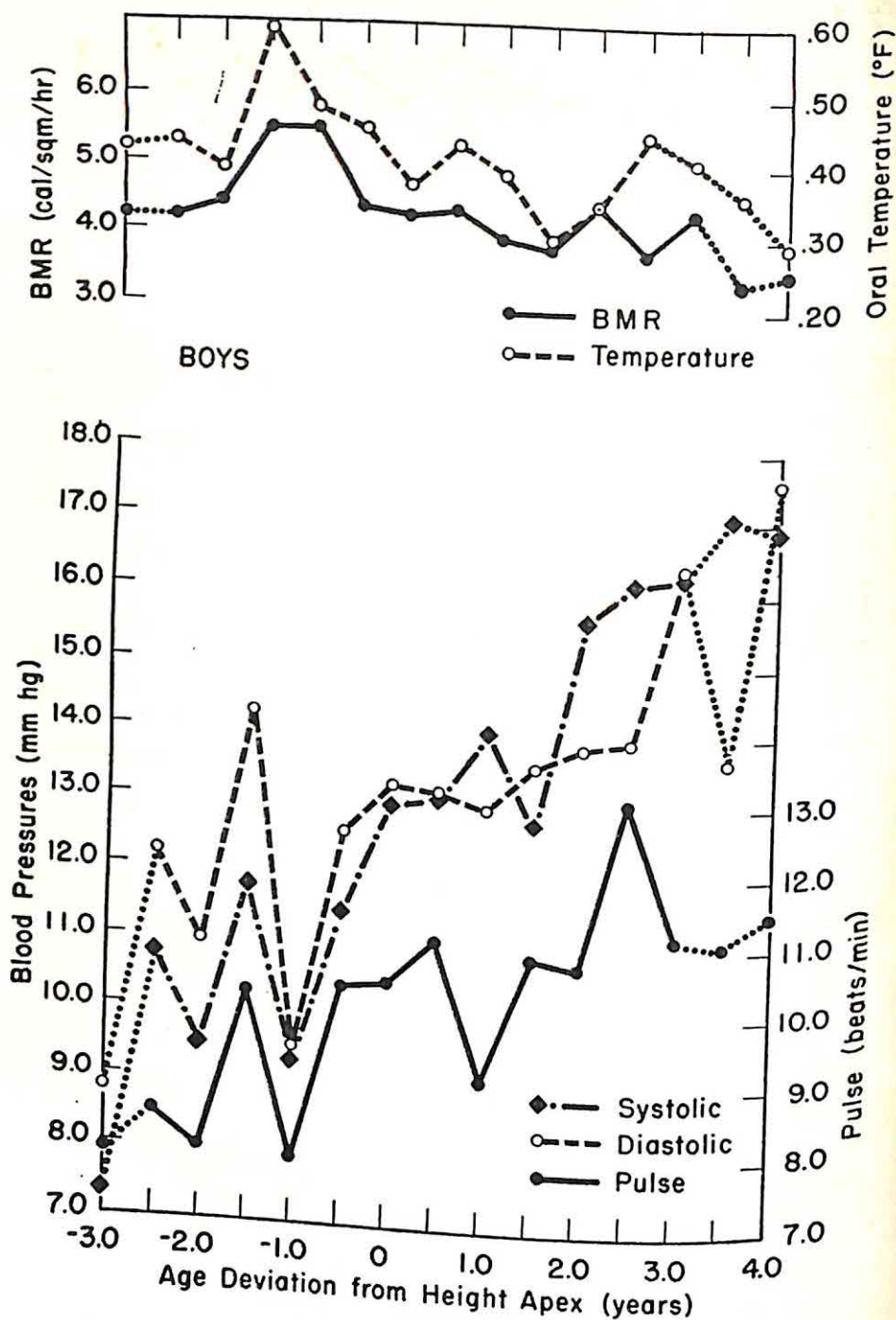


FIGURE 5—Mean range for basal metabolic rate, oral temperature, pulse rate, and systolic and diastolic blood pressure as a function of deviation from the age of maximum growth in height—boys.

tion examined. They certainly do not suggest a *general* increase in individual variability.

Relationship between intra-individual variability and "physiological age." Bayley (1), Jones (20), and Shock (28), among others, have demonstrated that for some purposes an index of "physiological age," or maturity, is more useful than chronological age. This should be especially true of physiological studies: individual differences in age and rate of maturing may obliterate systematic changes in variability which would appear if the data were grouped by degree of maturity rather than by chronological age. To examine this possibility we have plotted (Figures 5 and 6) the girls' data against menarcheal age and the boys' against deviation from the age of maximum growth in height (30). Since the previous plots (Figures 3 and 4) of diurnal variation were so unpromising, we give here only the data on the mean range.

We shall consider boys first. Figure 5 again shows the association between pulse and the blood pressures and between BMR and temperature. Furthermore, all five curves show a series of sharp inflections beginning two and a half years before the age of maximum growth. Since the curves are fairly irregular, it is possible that this rise and fall in all five curves is fortuitous. We are skeptical of its significance. Yet the fact remains that grouping the data according to the subjects' maturity rather than their chronological age should tend to eliminate fluctuations due to variations in apparatus and technique. That such grouping should align fluctuations for even a short part of the age span concerned strikes us as a bit surprising. It is a finding that should be tested with another sample of longitudinal measures.

What about girls? Figure 6 gives the mean ranges for girls plotted against menarcheal age. The curves take the same form as those for chronological age. Nothing in the figure suggests a temporary period of general increase or decrease in intra-individual physiological stability. The only menarcheal age at which all five measures increase at once is +3.0. This is so long after the menarche that we regard it as fortuitous, rather than as a genuine phenomenon associated with adolescence.

Summary

Physiological data from the files of the Adolescent Growth Study have been analyzed in terms of chronological age and in terms of physiological maturity in an effort to test the hypothesis that adolescence is characterized by a temporary decrease in physiological stability. For systolic and diastolic blood pressure and for pulse rate intra-individual variability increases throughout the period from 12 to 17 or 18 while variability in BMR decreases. Variability in body temperature shows little systematic change. Aside from these over-all slopes, no evidence of a temporary period of heightened variability was found for girls and only a slight indication was found for boys.

CHILD DEVELOPMENT

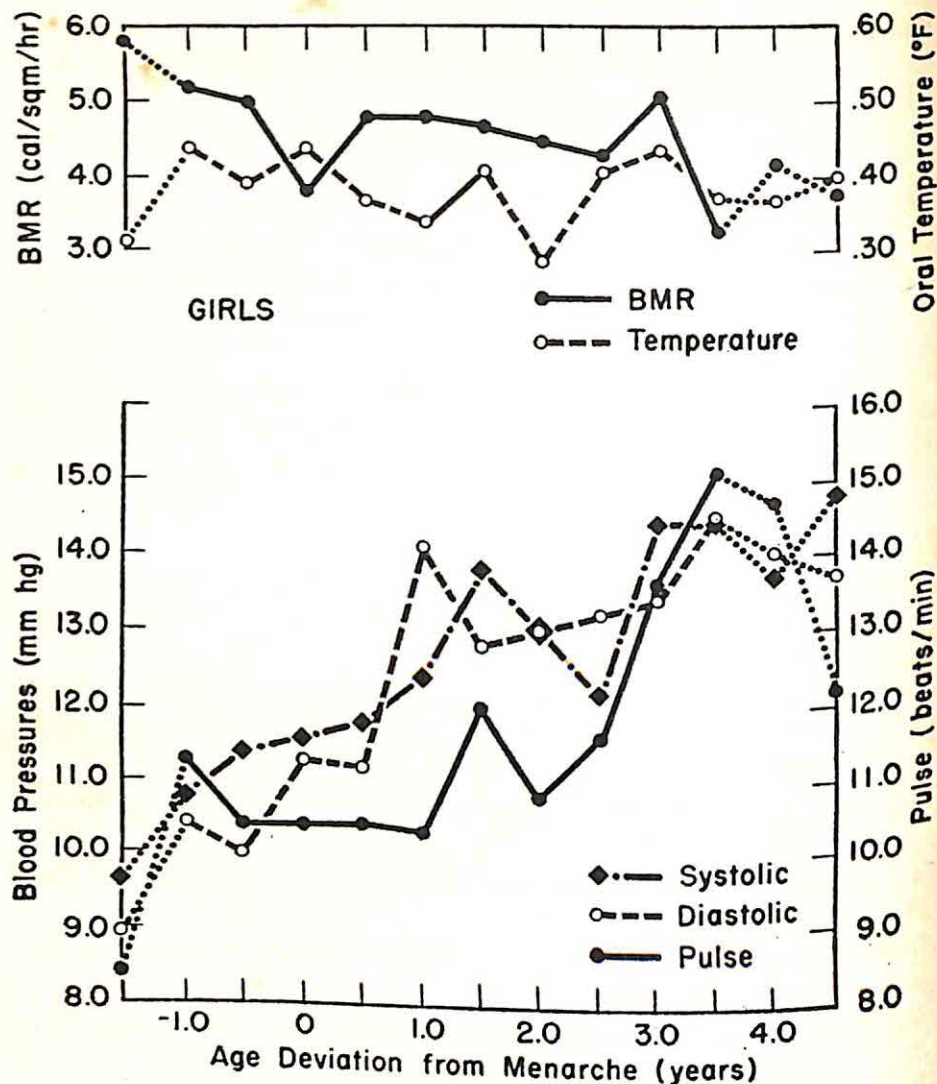


FIGURE 6—Mean range for basal metabolic rate, oral temperature, pulse rate, and systolic and diastolic blood pressure as a function of deviation from the age of menarche.

REFERENCES

1. BAYLEY, NANCY. Size and body build of adolescents in relation to rate of skeletal maturing. *Child Developm.*, 1943, 14, 47-90.
2. BAYLEY, NANCY, & STOLZ, H. Maturational changes in rectal temperatures of 61 infants from 1 to 36 months. *Child Developm.*, 1937, 8, 195-206.
3. BENEDICT, F. G. Factors affecting basal metabolism. *J. Biol. Chem.*, 1915, 20, 263-299.

4. BERKSON, J., & BOOTHBY, W. Studies of the energy of metabolism of normal individuals. The interindividual and intraindividual variability of basal metabolism. *Amer. J. Physiol.*, 1938, 121, 669-683.
5. BIERRING, E. *The standard metabolism of boys*. Copenhagen: Levin & Munsgaard, 1931.
6. BURLAGE, S. R. The blood pressure and heart rate, in girls, during adolescence. *Amer. J. Physiol.*, 1923, 64, 252-284.
7. EICHORN, DOROTHY H., & McKEE, J. P. Oral temperature and subcutaneous fat during adolescence. *Child Developm.*, 1953, 24, 235-247.
8. FREEMAN, G. L., & GIFFIN, L. L. The measurement of general reactivity under basal conditions. *J. gen. Psychol.*, 1939, 21, 63-72.
9. GRAHAM, A. W., HINES, E. A., & GAGE, R. P. Blood pressure in children between the ages of five and sixteen years. *Amer. J. Dis. Child.*, 1945, 69, 203-207.
10. GRIFFITH, F. R. Studies in human physiology: I. The metabolism and body temperature under basal conditions. *Amer. J. Physiol.*, 1929, 87, 602.
11. GROLLMAN, A. Physiological variations of cardiac output of man. VIII. The constancy of the cardiac output from day to day throughout one year. *Amer. J. Physiol.*, 1930, 93, 536-543.
12. HAFKESBRING, ROBERTA, & COLLETT, MARY. Day to day variations in basal metabolism of women. *Amer. J. Physiol.*, 1924, 70, 73-85.
13. HARMON, F. L. Reliability of metabolism measures by the closed circuit method. *J. appl. Physiol.*, 1953, 5, 773-778.
14. HARMON, F. L. Evaluation of oral temperature readings. *Science*, 1953, 118, 719-720.
15. HARRIS, J. A., & BENEDICT, F. G. The variation and the statistical constants of basal metabolism in men. *J. Biol. Chem.*, 1921, 46, 257-279.
16. HERRINGTON, L. P. The relation of physiological and social indices of activity level. In *Studies in personality*. New York: McGraw-Hill, 1942.
17. HORVATH, S. M., MENDUKE, H., & PIERSOL, G. M. Oral and rectal temperatures of man. *J. Amer. med. Ass.*, 1950, 144, 1562-1565.
18. JOHNSTON, J. A. Adolescence. In *Mitchell-Nelson textbook of pediatrics*. Philadelphia: W. B. Saunders, 1950. Pp. 1587-1600.
19. JONES, H. E. The adolescent growth study. *J. consult. Psychol.*, 1939, 3, 157-159; 177-180.
20. JONES, H. E. Motor performance and growth. *Univer. Calif. Publ. Child Developm.*, 1949, I.

CHILD DEVELOPMENT

21. KARVONEN, M. J., & NIEMI, M. The accuracy of basal metabolism determinations in the clinical laboratory. *Scand. J. Clin. Lab. Invest.*, 1952, 4, 112-114.
22. KLEITMAN, N., & RAMSAROOP, A. Periodicity in body temperature and heart rate. *Endocrinology*, 1948, 43, 1-20.
23. LUDWIG, GLADYS, N., & STOLZ, H. R. Temperature variations in normal pre-school children. *Child Developm.*, 1931, 2, 225-228.
24. PALMER, W. W., MEANS, J. H., & GAMBLE, J. L. Basal metabolism and creatinine elimination. *J. Biol. Chem.*, 1914, 19, 239-244.
25. PEELER, ANNIE L., HEPLER, OPAL E., KINNEY, VELMA M., CISLER, LILLIAN E., & JUNG, F. T. Normal values for serum cholesterol and basal metabolic rates and their correlation in normal man. *J. appl. Physiol.*, 1950, 3, 197-203.
26. RAGAN, C., & BORDLEY, J. The accuracy of clinical measurements of arterial blood pressure. *Bull. Johns Hopkins Hosp.*, 1941, 69, 504-528.
27. ROSENBLUTH, A. Der Grundunsatz bei Kindern und Jungendlichen. *Z. Kinderheilk.*, 1928, 46, 531.
28. SHOCK, N. W. The effect of menarche on basal physiological functions in girls. *Amer. J. Physiol.*, 1943, 139, 288-292.
29. SHOCK, N. W. Some physiological aspects of adolescence. *Texas Rep. Biol. Med.*, 1946, 4, 289-310.
30. STOLZ, H. R., & STOLZ, LOIS M. *Somatic development of adolescent boys*. New York: Macmillan, 1951.
31. SUTLIFF, W. D., & HOLT, EVELYN. The age curves of pulse rate under basal conditions. *Arch. intern. Med.*, 1925, 35, 224-241.
32. TALBOT, F. B., WILSON, E. B., & WORCESTER, J. Basal metabolism of girls. *Amer. J. Dis. Child.*, 1937, 53, 273-347.
33. WRIGHT, S. *Applied physiology* (9th Ed.). London: Oxford Univer. Press, 1952.

VARIABLES IN EARLY DISCRIMINATION LEARNING:

I. MOTOR RESPONSES IN THE TRAINING OF A LEFT-RIGHT DISCRIMINATION¹

WENDELL E. JEFFREY

University of California, Los Angeles

Recently there have been various discussions of perceptual development (4, 8), response factors in learning (7), or early discrimination learning (5) which are in general concerned with an organism's early experience with stimuli and the effects of this experience on subsequent performance. Experiments with both rats (3) and chimpanzees (9) have also demonstrated that early environment is important for later discrimination learning. For practical as well as theoretical purposes it would be valuable to know the type and amount of experience that is necessary, and the time at which it should occur, if the time factor is critical at all. Young children may be particularly good Ss for such research inasmuch as discriminations can be found which they do not normally make and seem to learn only with great difficulty before certain ages. The ability to learn these discriminations may then be taken as a critical indication of the effect of various types of experience in a pretraining situation.

A number of studies have been done (1, 6, 10) following such a plan but, in most cases, the problem presented to the child was simply one of learning which response to attach to each stimulus, inasmuch as the basic distinction among stimuli had probably already been learned.

For the present study a situation was sought where the presentation of differential stimuli apparently aroused no differential responses beyond those involved in receptor activation. This would provide a situation that would coincide more closely with Hebb's notion of early learning and would also fit closely with what others have referred to as perceptual development. A study by Swanson and Benton (11) suggested that left-right distinctions provide a difficult learning task for young children. After

¹ This research was supported by a grant from the research committee of the University of California and by research grant M-1498 from the National Institute of Mental Health, U. S. Public Health Service. I wish to express my appreciation to the staffs of the Small Fry and Peter Pan Nursery Schools for their cooperation in this research.

CHILD DEVELOPMENT

some preliminary research indicated that most four-year-olds were not capable of placing differential labels on stimuli that differed only in terms of left-right orientation, this task was chosen as a criterion discrimination.

It seemed possible that the ability to learn to respond discriminatively to stimuli differing in spatial orientation, i.e., left-right, might vary with the nature of the response to be learned. It was presumed that a response like pointing in the same direction in which stick figures were pointing would be easy to learn, and that the acquisition of this response would facilitate the acquisition of other responses. After preliminary research indicated that such a response was indeed easier than the labeling task, this experiment was set up to evaluate specifically the transfer effect from learning a discriminative motor response to the more difficult labeling task. It should be noted that, inasmuch as the direction of the motor response was not varied, no conclusion can be drawn regarding the importance of the symmetrical relationship of that response to the discriminative stimuli.

METHOD

Apparatus

The stimuli were two stick figures, with either a left or a right arm raised, each drawn in India ink on a 3 by 5 white card. A duplicate set of cards was available and was substituted during training to make certain that the children were not learning to discriminate dirt smudges or other non-experimental but consistent stimuli.

Two push buttons mounted 12 inches apart on top of a box provided the motor response. Tape recorded children's music through earphones served as incentive and reinforcement. The music was controlled by a foot button and Hunter timer, so that for a correct motor or verbal response *E* pushed the foot button permitting *S* to listen to 10 seconds of the recorded material.

Procedure

The experiment followed what was essentially an X-Y-X transfer design in which the X task involved training to label the stick figures differentially, and the Y task trained *S* to make button pressing responses in the general direction toward which the figure could be said to be pointing. The control group received only the X, or labeling, training.

Each *S* was asked to come into an adjoining room to play a game. At this time, he was introduced to the earphones and seated in a chair in front of the box. Typically there was some small talk about earphones and airplane pilots. Then, to help accustom the child to the situation, the music was turned on and allowed to play for a minute or two.

To start the initial training on the criterion, or X, task, the music was stopped and the stimuli were placed side by side on the box in front of *S*. He was then asked if he could tell how the two figures differed. None of

the Ss gave any evidence of seeing the difference or of knowing what was meant by the word, "difference." The differences were then pointed out, and S was told that the figure with the arm pointing "in this way" would be called "Jack," and the figure with the arm pointing "this way" would be called "Jill," and that the music would come on only if he gave the right name for each picture whenever the picture was presented. The stimuli were presented successively following a randomized order.

For the control group, training proceeded as described above until S met a criterion of 10 successive correct responses, or until he quit and refused to return, or until such time as E decided to discontinue training. The latter was done only when there was no evidence of learning after at least 80 trials. At intervals from 10 to 20 trials, instructions were given again, pointing out the relationship between the names and directions in which the arms were pointing.

The experimental group first received 20 trials of the same training as the control group with instructions repeated after the 10th trial. After this initial training on the control task, the experimental Ss were instructed to push the button toward which the stick figure was pointing, and the appropriate button for each stick figure was indicated by E. The experimental Ss were trained on this task to a criterion of 10 successive correct responses and then returned to the labeling task. Two of the nine children who were given the experimental treatment went back to the original task in the same session. Seven Ss came back on a second day one week later, and were given 10 more button pressing trials before returning to the labeling task.

Subjects

Twenty-eight children were divided into two groups, matched for mean CA of four years and four months and range from three years, eleven months, to four years, nine months. The training procedure was assigned to each S by the toss of a coin. There were seven girls and seven boys in each group.

RESULTS

A contingency table showing the number of Ss in the experimental and control groups who were able to meet the criterion of 10 successive correct discriminations in the labeling task is presented in Table 1. The Fisher exact probability test, which tests the hypothesis as to whether two groups differ significantly in the proportion of passes and fails attributed to them, indicates that the null hypothesis can be rejected at better than the .005 level of confidence.

However, it should be noted that three Ss in the control group were able to perform the labeling task without the experimental treatment. The control Ss can be categorized roughly into three groups: first, three who performed at a high level from the beginning and completed the criterion

CHILD DEVELOPMENT

TABLE 1

CONTINGENCY TABLE SHOWING NUMBERS IN CONTROL AND
EXPERIMENTAL GROUPS PASSING AND FAILING THE
LABELING TASK

	<i>Pass</i>	<i>Fail</i>
Experimental	13	1
Control	3	11

within 30 to 40 trials; second, four who refused to continue after rather limited number of trials and five who refused to return after low success in the initial session; third, two who continued to come back with no evidence of ever responding above chance.

Inasmuch as the labeling task proved to be less difficult than was indicated by other studies (2, 11), it seemed important to try to identify any members of the experimental group who might learn the labeling task without the experimental treatment, in order to eliminate them from consideration in the final results. Thus, two *Ss* in the experimental group were given an extra 20 trials on the initial test task because of high performance on the first 20 trials. One of these met the criterion while the other did not, even though the instructions had been repeated before the 21st and 31st trials. Four other children in the experimental group also showed evidence of learning or met the criterion task during the initial 20 training trials.

Table 2 gives a picture of the results if one removes from Table 1 the three *Ss* of the control group and the five *Ss* of the experimental group who learned the criterion task so readily that it seems reasonable to assume that they were not examples of early or original learners. This table gives a more striking picture of what has taken place. Again, the null hypothesis can be rejected at the .005 level of confidence. As can be seen, those *Ss* in the control group who did not show promise of learning from very early

TABLE 2

CONTINGENCY TABLE OF NUMBERS PASSING AND FAILING
THE LABELING TASK WHEN CHILDREN WHO APPARENTLY
LEARNED WITH EASE WERE REJECTED FROM
BOTH GROUPS

	<i>Pass</i>	<i>Fail</i>
Experimental	8	1
Control	0	11

in the training did not learn subsequently. In the experimental group, when those who apparently were capable of learning the criterion task with some ease were excluded from the group, thus making this group comparable to the non-learning controls, we still found that all but one of these *Ss* learned the criterion task quickly after the experimental treatment. The one *S* who did not learn was performing consistently above chance (eight out of 10, and nine out of 10 in the last 20 trials) but was apparently too distractable to meet the criterion of 10 successive correct responses. It should be noted that in the experimental treatment four *Ss* performed perfectly after instructions alone, four *Ss* completed the criterion in 20 trials, and one took 30 trials. Subsequent to this training, seven *Ss* completed the criterion on the test task within 20 trials, one completed the criterion within 30 trials, and, as mentioned above, one failed to meet the criterion within 50 trials and training was discontinued.

DISCUSSION

The exploratory nature of this research makes it hazardous to discuss it on a theoretical level. The left-right discrimination is really somewhat easier than had been expected. Thus, it is probably gratuitous to suggest that we are getting at early learning in the sense that we had hoped. We are now working with pitch discrimination which appears to be almost universally difficult and better meets our criterion for early learning. From this experiment, however, it seems reasonable to conclude that some discriminations that may be difficult for children to make can be simplified so that the child can learn the simpler task, and that, as the result of learning the simpler discrimination, performance may be facilitated in more complex situations.

Simplification can be specified in both the stimulus and response dimensions. In this experiment the response was simplified. What will happen if the stimulus dimension is manipulated remains to be demonstrated. It may be that exaggerating the differences between stimuli in order to simplify them may have the net effect of eliciting a response more easily. Thus, in the present experiment lengthening the arms of the stick figures might have increased the number of successful performances with the naming task by producing in the observing response the sort of distinct stimulation now produced by the button pressing response. It should be reiterated that the button pressing response apparently did more than simply call attention to the appropriate part of the stimuli inasmuch as the importance of the direction that the arms were pointing was demonstrated to *S* several times during the test trials.

Whether the button pressing served a mediating function was not clear. Preliminary research most definitely indicated that it was not necessary to set up a series of trials attaching names to the buttons. A number of the children, however, obviously used the button pressing response or fractional

CHILD DEVELOPMENT

parts of it in learning to apply the names. In one case, a child who had to be told twice to discontinue the button pressing started lifting the appropriate shoulder for each stimulus before applying labels. Other children would look at the appropriate button before supplying the name of the figure. The majority, however, displayed no obvious remnant of the motor response when learning the names. Further research manipulating both stimulus and response variables is obviously a necessity before one speculates further about the basic processes involved or about other levels of mediation.

At a practical level, these data can be interpreted as having fairly direct implications with regard to specific instruction, such as teaching of reading. Research by Davidson (2) has indicated that children are not normally capable of discriminating "b" from "d" until seven and one-half years of age. The present data indicate that a child incapable of performing this discrimination at a given time can be taught it very quickly, even as early as four years of age, if conditions are properly arranged.

The recommendation to delay training after the failure of children to meet certain readiness criteria probably arises from observations of the frustration produced when children attempt to learn discriminations that are too difficult for them, rather than from knowledge of the processes involved in discrimination learning. The refusal of a number of the children in this experiment to remain in the situation under conditions providing 50 per cent reinforcement was striking when compared with the extended periods of time others spent if success was relatively high. Four-year-old children have already learned to label such performance as failure and react by withdrawing from the situation. However, it should be noted that the situation in general was sufficiently reinforcing, if success was relatively high, to keep children in the experimental situation for extended periods of time and also to keep them coming back on subsequent days. Furthermore, these children were from schools which attempted to minimize the importance of high performance and competition. Thus, it undoubtedly is wise not to push the child when he has difficulty. However, the present study suggests that one can continue training at a different level and thus achieve the desired results without concomitant frustration.

SUMMARY

An experiment was devised to test the facilitative effect of attaching simple motor responses to stimuli differing in spatial (left-right) orientation on the subsequent labeling of these stimuli. Four-year-old children were used as Ss and it was found that, except for a few children who learned the labeling task with ease, the majority of four-year-olds found the task essentially impossible.

The experimental treatment of learning to press buttons oriented in the directions the stick figures used as stimuli were pointing was readily learned

by all Ss in the experimental group and this training was found to have a significant effect on subsequently learning to apply labels.

REFERENCES

1. CANTOR, G. N. Effects of three types of pretraining on discrimination learning in preschool children. *J. exp. Psychol.*, 1955, 49, 339-342.
2. DAVIDSON, H. P. A study of the confusing letters, d, b, p, and q. *J. genet. Psychol.*, 1935, 47, 458-468.
3. FORGUS, R. H. Early visual and motor experience as determiners of complex maze-learning ability under rich and reduced stimulation. *J. comp. physiol. Psychol.*, 1955, 48, 215-220.
4. GIBSON, J. J., & GIBSON, E. J. Perceptual learning—differentiation or enrichment? *Psychol. Rev.*, 1955, 62, 32-41.
5. HEBB, D. O. *Organization of behavior*. New York: Wiley, 1949.
6. JEFFREY, W. E. The effects of verbal and nonverbal responses in mediating an instrumental act. *J. exp. Psychol.*, 1953, 45, 327-333.
7. MANDLER, G. M. Response factors in human learning. *Psychol. Rev.*, 1954, 61, 235-244.
8. POSTMAN, L. Association theory and perceptual learning. *Psychol. Rev.*, 1955, 62, 438-446.
9. RIESEN, A. H. The development of visual perception in man and chimpanzee. *Science*, 1947, 106, 107-108.
10. SPIKER, C. C., & TERRELL, G. Factors associated with transposition behavior of preschool children. *J. genet. Psychol.*, 1955, 86, 143-158.
11. SWANSON, R., & BENTON, A. L. Some aspects of the genetic development of right-left discrimination. *Child Developm.*, 1955, 29, 123-133.

RORSCHACH BEHAVIOR AND PERFORMANCE OF HIGH AND LOW ANXIOUS CHILDREN¹

SEYMOUR B. SARASON, KENNETH DAVIDSON, FREDERICK
LIGHTHALL, *and* RICHARD WAITE
Yale University

The present article deals with the performance and behavior of high anxious (HA) and low anxious children (LA) in the Rorschach situation. This study is part of an ongoing research project on the measurement and correlates of anxiety in school age children. Previous reports (3, 6, 7, 8, 9) from this project have described the construction of Test Anxiety (TA) and General Anxiety (GA) Scales and, in addition, presented data indicating that these scales had an encouraging degree of validity.

In an attempt to explore further the validity of these scales we chose the Rorschach situation because it is one in which the child has to solve problems involving culturally unfamiliar stimuli with practically no aid or direction from another person. The method of administration with its underlying rationale has previously been described by Sarason (5, p. 111). It should be emphasized that the method of administration involves no questioning or prodding in the performance and very little questioning in the inquiry. Briefly, our expectations were that in a problem-solving situation in which the child is required to make his own decisions, there would be more signs of interference and ineffectiveness in the functioning of the HA child than in the case of the LA child.

SUBJECTS

Thirty-two pairs of subjects were matched for grade, sex, and average verbal and nonverbal T-score on the Otis Quick-Scoring Mental Ability Test (Alpha). They differed in that one member of each pair was in the

¹ The project of which this study was a part is being supported by a grant from U.S.P.H.S. We wish to express our deep appreciation to Iris B. Keim and Joan Bliss for their aid in this study.

fourth quartile of scores on Test Anxiety and General Anxiety Scales (the HA subject) while the other member of the pair was in the first quartile (the LA subject). There were 16 boys and 16 girls in each of the two anxiety groups. The 32 pairs were drawn from a sample of 747 subjects in grades 1 through 4, who had been given the anxiety questionnaires several months earlier. At the time of the Rorschach the subjects were distributed in grades 2 through 5. The total sample of 747 excluded subjects with a known academic or behavior problem and those whose parents were divorced or separated.

PROCEDURE

The Rorschach testing was done by two female examiners. Following the inquiry of each response, the subject was given a piece of tracing paper with the request to trace the response so that the examiner "could see it just as you did." After the test was over, a description of the subject's behavior was written, averaging approximately one and one-half double-spaced pages in length. These behavior descriptions were not made with any fixed schedule of traits in mind. The examiners were told to describe what to them were the dominant characteristics of the child's behavior, and also to include their own spontaneous reactions to the child's behavior.

One of the writers (S.B.S.) spent approximately 20 to 40 minutes in a somewhat superficial clinical analysis of each protocol (including the behavior description). Brief notes were made during the analysis, and in each case a judgment was made as to whether the subject was HA or LA, i.e., whether the child answered the questionnaires as an HA or LA subject. Following this, the 64 records were arranged for the clinician into the 32 matched pairs with the purpose now of re-evaluating the previously recorded judgments and making whatever changes seemed indicated when examining each member of the pair together, the group to which each member belonged still being unknown to the clinician. The comparison of the matched pairs was done before the results of the first analysis were known.

The clinical analyses were based both on theoretical conceptions concerning anxiety as well as the particular psychologist's clinical experience with children. Therefore, the major criteria employed subjectively by the clinician were then explicitly stated and tested statistically.

The protocols were also scored in the conventional manner, using Hertz's (4) manual as a guide. The two Rorschach examiners independently scored 10 protocols and agreement on 73 per cent of the responses was obtained. Each of the remaining 54 protocols was then scored by both examiners in conference. Where there was a difference in scoring a response, an additional judge was consulted. It might be helpful to record our opinion that, in the absence of tracings, the scoring of these records would have been markedly less reliable than it was.

In the present study, as in all other aspects of this project, the collection and analysis of the data were done without knowledge of the group to which a child belonged.

RESULTS

Clinical Analysis of the 64 Records

Forty of the 64 cases were placed in the proper anxiety group, the p value for this result being .03. When the 64 records were then arranged into the 32 matched pairs with the opportunity to make new judgments, 18 of the 32 pairs were correctly categorized, this result being significant at only the .30 level. The clinician making the judgments experienced great difficulty with 12 of the matched pairs, these pairs containing cases who in the previous analysis had also been a source of difficulty. In making his final judgment with these 12 pairs the clinician explicitly indicated that he was making his judgments with much doubt. *The doubt did not concern the judgment about the subject's anxiety in the Rorschach situation but whether or not the subject would admit to anxiety on the questionnaires.* It should also be noted that all judgments on the non-questioned pairs took approximately 30 to 45 minutes. In the 20 pairs which were not questioned, 15 were correct ($p = .03$). (The above p values were obtained using the sign test procedure.)

When these results were known, the clinician then reviewed the incorrectly matched pairs in order to determine, if possible, the reasons for the mismatching. It quickly became clear that there were a number of individuals in these pairs for whom the examiner explicitly had used such phrases as: "He is a very guarded child," "I felt I had no rapport at all with this child," "He was very defensive and I don't know what was going on inside." It also seemed as if it was largely LA cases who were so described and that such behavior was interpreted by the clinician as a reflection of anxiety in the child. The two examiners independently then read each of the 64 behavior descriptions with the instruction to put each child into one of two categories: The child was clearly of the "guarded" type or he was not. There was complete agreement in 13 cases that the child was guarded. There were six cases where one examiner categorized the child as guarded while the other examiner had not, although the latter judge explicitly questioned her own categorization in these six instances. There was no instance where one examiner judged a child as guarded while the other examiner clearly did not.

The 13 cases in which there was complete agreement were largely LA (nine LA vs. four HA) and included seven children who were members of mismatched cases. Of the six cases where there was partial agreement about the child's guardedness, four were HA and two LA, with only one a member of a mismatched pair. In other words, of the 12 pairs which the clinician had mismatched and about which he had felt great uncertainty in judging, eight contained a child who was guarded or defensive or with

whom the examiner could not establish rapport. Six of the eight were LA cases labelled HA by the clinician. It would seem, therefore, that one of the sources of error in matching was the child who admitted to little or no anxiety on the two scales but whose guarded behavior in the Rorschach situation was interpreted by the clinician as symptomatic of anxiety. From the inception of this project we had assumed that, if a child received high scores on both scales, we could place greater credence on his report than on that of the child who admitted to little or no anxiety. The analysis of the mismatched supports this assumption.

Criteria Employed by Clinician

What follows here concerns those criteria which the interpreting clinician employed and could state in a way so as to make statistical testing possible. The interpretive process is too complex and too little understood and studied to expect that, at this point at least, one can do other than evaluate aspects of the process. It is easy for a clinician to say that he employed a criterion, but it is extremely difficult for him to indicate how he weights it when other criteria are present in varying degrees.

Rejection of cards. The inability to respond is perhaps the most blatant indicator of interference in response. In the HA group 13 of the 32 cases rejected at least one card while in the LA group there were six such cases. A chi square analysis resulted in a p value of .05. (This and all subsequent p values are based on a one-tail test.)

Minus response (F-/R). This measure reflects the degree to which an individual's responses do not correspond to the stimulus area employed. In a chi square analysis in which the cut-off point was the median of the F-% distribution, significantly more HA were above the median ($p = .025$). When the vague responses were taken into account the differences between the groups disappeared, contrary to the clinician's expectations. In other words, to the extent that the clinician attributed the same weight to giving many vague responses as to a high F-%, he was reducing his valid discriminations between the two anxiety groups.

Total number of responses. The expectation that anxiety would interfere with output is, of course, similar to the expectation concerning card rejections. In a chi square analysis there was a tendency for HA subjects to give fewer responses ($p = .05$). When cases in the two anxiety groups who rejected cards were excluded from the analysis, the probability level remains the same.

Anatomy responses. It is assumed that the high anxious individual is one who, consciously or unconsciously, is or has been concerned with body adequacy or integrity, and that such concern will be reflected in the content of responses. The following are examples of anatomy responses: skeleton, inside of somebody, X-ray, lungs, person's breast, tonsils, etc. Thirteen HA individuals gave at least one anatomy response while six in the LA group did so ($p = .05$). For reasons similar to those above, it was

expected that more HA than LA individuals would give responses in which somebody or something was explicitly killed or damaged. There was no difference at all between the groups in this type of content.

Active aggressive responses. It is assumed that the anxious individual tends to perceive himself as inadequate and has difficulty justifying the direct expression of aggressive feeling (3). On this basis it was expected that more LA than HA individuals would give responses involving fighting, arguing, or a "volcano" percept. A chi square analysis indicated that more LA than HA individuals gave such responses ($p = .05$).

The above results suggest that most of the criteria which the clinician could explicitly state had discriminatory value.

Results with Conventional Scores

More HA than LA individuals were unable to respond at all to any objective property of the cards involving color (chromatic, achromatic, shading), the p value being .025. Seventeen of the 32 HA cases could not respond to any aspect of color while eight in the LA group could not do so. This finding, which was the only one of the conventional scores which attained or was near significance, can be integrated with some of the findings in the previous section in the following way: in a problem-solving situation containing relatively unfamiliar stimuli to which he must respond independently (i.e., deciding for oneself how, when, and how often to respond) the HA individual will, when he can respond at all, tend to reflect in his responsiveness illogical or irrational ways of thinking, and such responsiveness will tend not to incorporate the obvious properties of the stimulus. This conclusion, which is akin to one by Cox and Sarason (2) in a similar study of college students, suggests that one of the effects of anxiety is to exacerbate to an interfering degree the role of internal and affective factors in outwardly directed responsiveness.

Analysis of Behavior Notes

The first analysis involved a count of those words (e.g., afraid, uneasy, shy, disturbed, etc.) which might be symptomatic of anxiety. Some of these words appeared with too little frequency to warrant any conclusion. For any single word category there were no notable differences between HA and LA groups. When a count was made of the number of different categories in which an individual was entered, there was a tendency for HA boys to be entered in more categories than was the case for LA boys.² There was no such difference between HA and LA girls. The only other tendencies which seemed worthy of note concerned HA boys and HA girls: eight or 50 per cent of the HA girls were labeled as "nervous" by the examiner whereas only two of the HA boys were so designated; 11 or 65 per cent of the HA boys were described as uncertain or unsure of

² In the analyses previously presented there were no sex differences which were significant or approached significance.

themselves in this situation, whereas only three HA girls were so described. No such differences were found between LA boys and girls. It may be that a word like "nervous," on the one hand, and words like "unsure" and "uncertain," on the other hand, did not in fact signify different kinds of behavior. However, it is our impression from the context in which these words appeared that "nervous" signified a general impression of the examiner (e.g., "a nervous type of child," "a nervous, high strung child") whereas "unsure" and "uncertain" seemed to refer more specifically to reactions to the problem-solving task.

In order to check further on the behavioral differences suggested above, a count was made of the number of questions which the child asked the examiner during the performance part of the Rorschach, that part during which the examiner was most nondirective and the factor of unfamiliarity of the stimulus task presumably most strong. We were interested in such a count because we assumed that the asking of questions reflected not only dependence but uncertainty as well. It was found that three HA girls asked questions whereas nine HA boys did so. This suggests that the label "nervous" in connection with HA girls may well reflect behavior different from that signified by "uncertainty and unsureness" in HA boys. In the other anxiety group six LA girls and three LA boys asked questions. Again we found that the differences between HA and LA boys were greater than between LA and HA girls.

A further analysis, relevant to the above, concerned the child's handling of the tracing of each of his responses. The focus here was whether a child said he could not or did not know how to trace a response, or stated that he was not satisfied with what he had traced, or where the examiner explicitly stated that the child was concerned in some way with having to trace his responses. As in the previous analyses it was the HA boys who experienced difficulty with tracings: nine HA boys vs. four HA girls, and four LA boys vs. five LA girls. In going over the count it became apparent that some children had been described by the examiner as being concerned either if in tracing they tended to check their tracing with the blot or if they took back the tracing from the examiner in order to add something that had been forgotten. Since such behavior seems much less clearly indicative of concern than a child's own statement of concern, a similar count was made excluding these cases. In the new count there were eight HA boys vs. two HA girls, and two LA boys vs. four LA girls. The tendency for differences to appear between HA and LA boys, but not between HA and LA girls, remained.

DISCUSSION

We have already indicated that we chose the Rorschach situation because it is one in which the child had to decide for himself how to handle a problem-solving task containing relatively unfamiliar stimuli. We might

put our intention in another way. We were attempting to increase the possibility that a child would experience a situation as one of danger in that external support for the handling and fulfillment of his needs would be minimized or nonexistent. Because it is in such a situation that anxiety should be experienced, it was predicted that an HA group would be more adversely affected than would the LA group. We evaluate the findings we have presented as supporting such a prediction. It seems appropriate to indicate at this point that the findings of the present study are similar to those obtained with HA and LA college students, despite the obvious differences in the nature of the two samples (2).

One of the most thorny problems in the development of truly discriminating personality questionnaires, especially when the subject is asked to reveal what may be termed weaknesses, is that the subject can wittingly or unwittingly give an invalid response. The findings in the present study in relation to the "guarded" child underlines the importance of the problem but throws relatively little light on two questions. How can one pick out *from questionnaires* the child who admits to little or no anxiety on the questionnaire but whose overt behavior in relevant situations suggests the experience of anxiety? How can one discriminate between the child who consciously gives an invalid low score and the one whose perception and remembrance of his test behavior is subject to defensive distortion? In an attempt to explore this problem a "lie" scale consisting of 11 items was developed and embedded in the general anxiety scale. Each of these 11 items was specifically concerned with the lie tendency in relation to anxiety or worry (e.g., "I have never had a scary dream," "Do you ever worry?"). The correlation between lie and anxiety scores tended to be of such size (ranging from $-.65$ to $-.75$) that one could not assume that the two scores were measuring different things. This finding is markedly at variance from that of Castenada, McCandless, and Palermo (1) who found practically no correlation between anxiety and lie scores. It is important to point out that whereas our lie items specifically concerned anxiety their items were more omnibus in nature. It is our opinion that it is erroneous to assume that the lie or defensive tendency is not dependent on the content area (e.g., anxiety, aggression, cheating) about which one is questioning the subject. If the lie items concern cheating or kindness, and the scale in which they are embedded concerns anxiety, there is no compelling *a priori* reason to expect a significant correlation between the lie and anxiety scales. It would seem that until more thought is given to the theoretical rationale and methodology of the measurement by questionnaire of different kinds of defensive tendencies the clinical and research utility of personality questionnaires will be limited.

An important problem suggested by this study stems from the suggestive trend in the behavior notes, namely that the overt behavior of HA boys seems to reflect anxiety differently from that of HA girls. In addition, the overt behavior of HA boys seemed different from that of LA boys

CHILD DEVELOPMENT

whereas no such trend was discernible between HA and LA girls. Although these trends may be unreliable because of the small number of cases upon which they are based, we presented the behavioral data because of their possible significance for other of our findings. Twenty-four of the 32 matched pairs of subjects were also utilized in a learning study (8). In that study significant differences between HA and LA subjects were found but the difference between the HA and LA boys was notably larger than that between HA and LA girls. In addition, although girls rather consistently get significantly higher anxiety scores than boys, the correlation between anxiety score and score on conventional group tests of intelligence is no higher among girls than among boys. One possible implication of these different findings is that a high anxiety score does not have the same psychological significance in a boy as in a girl. We have elsewhere (7) advanced the hypothesis that our culture makes it much more difficult for a boy than for a girl to admit explicitly to anxiety (or weakness). For a girl to admit to anxiety is not as likely to impair her own or others' evaluation of her femininity as a similar admission in a boy would impair his own or others' evaluation of his masculinity. If this is correct, one might expect that the anxiety which the HA girl admits to on the questionnaires does not have as interfering an effect on her problem-solving behavior as in the case of the HA boys. The fact that boys get lower anxiety scores than girls—which we interpret as a reflection of a kind of learned defensiveness or suppression in boys—would suggest that the boy who does get a high score is one who has difficulty handling his anxiety. The finding in the present study concerning the dependent behavior of HA boys fits in with such a hypothesis.

SUMMARY

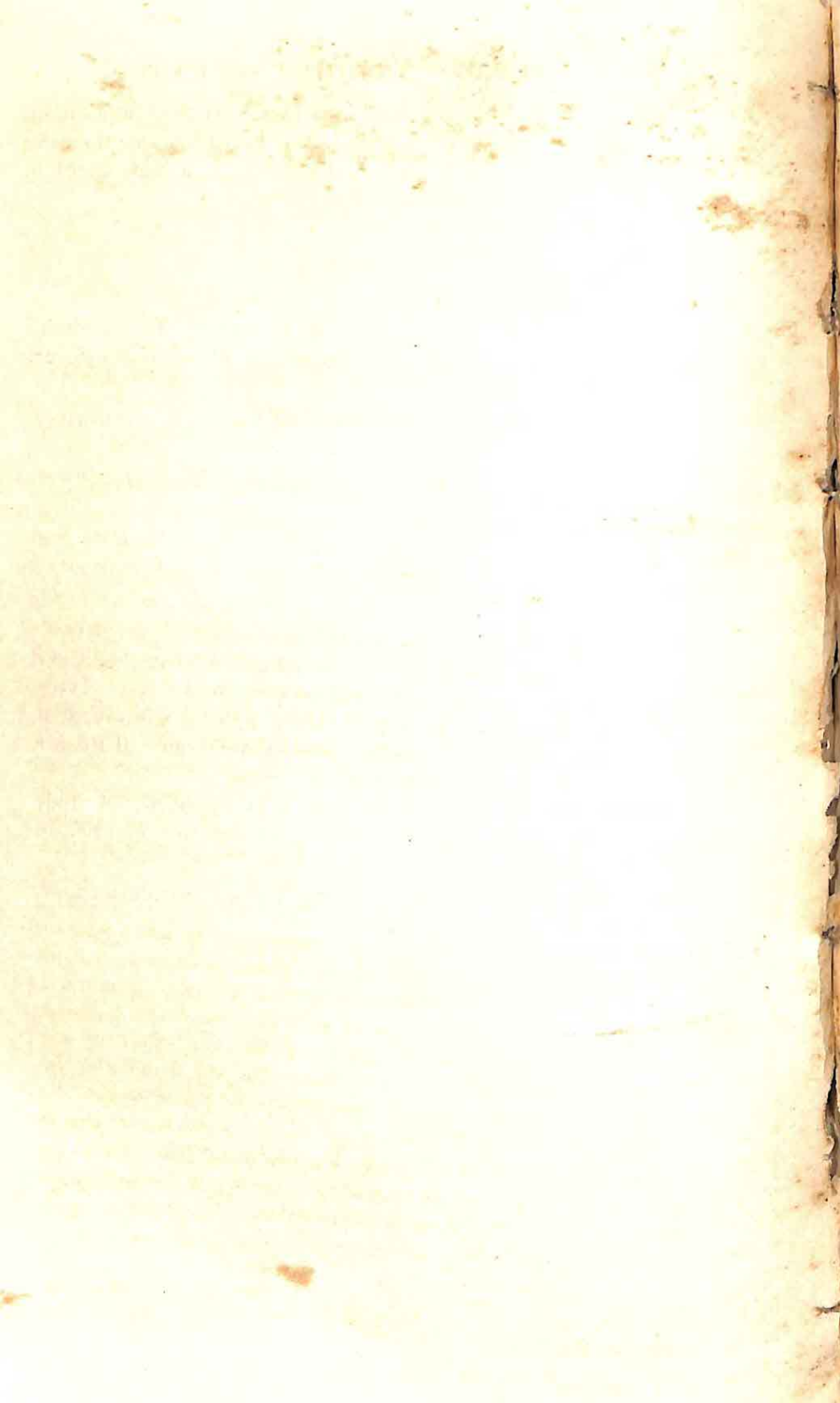
The Rorschach was administered to 32 high and 32 low anxiety children who were matched for grade, sex, and IQ score. A relatively brief and blind clinical analysis (case-by-case) by a clinician resulted in significant discriminations between the two anxiety groups. In a matched-pair analysis the discrimination was not significant. Further analysis revealed that one of the sources of error in matching was the child who admitted to little or no anxiety on the anxiety scales but whose guarded behavior in the Rorschach situation was interpreted by the clinician as symptomatic of anxiety.

In contrast to the low anxious subjects, the high anxious rejected more cards, gave fewer responses, gave fewer responses with aggressive content, gave more responses with anatomy content, and responded less to any aspect of color (chromatic, achromatic, shading). An analysis of the behavior notes suggested that the HA boys are more easily differentiated from the LA boys than are the HA girls from the LA girls. The HA boys present a rather clear behavioral picture of indecision and dependence.

The results of this study are discussed in terms of (a) their implications for the measurement of defensive tendencies as they affect questionnaire taking and (b) the possibility that identical scores (e.g., a high score) in boys and girls have different psychological significances.

REFERENCES

1. CASTANEDA, A., MCCANDLESS, B. R., & PALERMO, D. S. The children's form of the manifest anxiety scale. *Child Developm.*, 1956, 27, 317-326.
2. COX, N., & SARASON, S. B. Test anxiety and Rorschach performance. *J. abnorm. soc. Psychol.*, 1954, 49, 371-377.
3. DORIS, J. Test anxiety and blame assignment in children. *J. abnorm. soc. Psychol.*, in press.
4. HERTZ, M. *Frequency tables for scoring responses to the Rorschach Inkblot Test*. Cleveland: Western Reserve Univer. Press, 1951.
5. SARASON, S. B. *The clinical interaction: with special reference to the Rorschach*. New York: Harper, 1954.
6. SARASON, S. B., DAVIDSON, K., LIGHTHALL, F., & WAITE, R. A test anxiety scale for children. *Child Developm.*, 1958, 29, 105-113.
7. SARNOFF, I., LIGHTHALL, F., WAITE, R., DAVIDSON, K., & SARASON, S. B. A cross-cultural study of anxiety amongst American and English school children. *J. educ. Psychol.*, in press.
8. WAITE, R., SARASON, S. B., LIGHTHALL, F., & DAVIDSON, K. A study of anxiety and learning in children. *J. abnorm. soc. Psychol.*, in press.
9. ZWEIBELSON, I. Test anxiety and intelligence test performance. *J. consult. Psychol.*, 1956, 20, 479-481.



CLASSROOM OBSERVATIONS OF HIGH AND LOW ANXIOUS CHILDREN¹

SEYMOUR B. SARASON, KENNETH DAVIDSON, FREDERICK
LIGHTHALL, and RICHARD WAITE
Yale University

This study is part of an ongoing project concerned with the development and correlates of anxiety in children of school age. Specifically, the present study involves the direct classroom observations of high (HA) and low (LA) anxious children, anxiety being defined here by high scores both on a test anxiety scale and a general anxiety scale previously described (1, 3, 4). In the initial phases of this project the focus was on the development of the two scales and the establishment of their face validity. In the present phase the focus has been on the study of 32 matched pairs of HA and LA children in the following ways: an experimental learning situation (5), the Rorschach (2), parental interviews (1), and direct observations in the classroom.

A number of different problems were encountered in attempting to develop a methodology for observations in the classroom. In a pretest both an HA and an LA child in the same classroom were observed for approximately three hours by one observer. This was done for eight pairs of children. One child was observed for 15 minutes, the other child for the next 15 minutes, and so on until the noon recess. The aim of the observation was to record as much of the child's behavior as was possible. We assumed that such an aim was presumptuous but we also assumed that it would be better in a pretest to err in getting too many than in getting too few data. As it turned out, however, we encountered a surprisingly restricted range of behavior, the class usually being structured in such a way that there were few times when a child had an opportunity to

¹ The project of which this study was a part is being supported by a grant from U.S.P.H.S. We wish to express our appreciation to Iris Keim and Joan Bliss for their help in collecting the data.

CHILD DEVELOPMENT

engage spontaneously in interpersonal contact. Most of the time the child was engaged in directed activity, such as silent reading, preparation of lessons, or attending to what the teacher was doing or saying. There were a few times during each observation when there was interaction between the observed child and the teacher (or another child) the content of which was ambiguous because the observer could not hear or accurately see what was going on.

A limitation of the blow-by-blow type of observation is that it is extremely difficult for someone who reads a protocol to get something which resembles an integrated personality picture. It was our experience that although the observer could integrate the many details into a distinctive personality picture, another reader could not do so because he, unlike the observer, did not have a reservoir of impressions to draw on, i.e., the detailed notes did not have the surplus meanings for the reader that they had for the observer.

We at first thought that our problem would have been much more simple if we had a clearer idea of the kinds of overt behaviors which could and should be observed in the classroom observations of HA and LA children. This was undoubtedly true, but our experiences have indicated that even if we were to begin with well-defined, theoretically-based categories of behavior, one would have to carry out the observations over a relatively long period of time before obtaining frequencies which would justify statistical analysis. The importance of observations extended over time is further emphasized by the variability of a child's behavior. During the course of the pretest as well as the main study we found children whose behavior varied markedly from morning to afternoon, from one day to the following day, and in a few instances, unaccountable changes in behavior within a single observational period. Aside from these presumably intraindividual variations we also encountered the situation where a child's behavior on a second observation was different from the first because of an unusual external occurrence, e.g., the regular teacher was absent, or a remedial specialist took over the class.

Because the 32 matched pairs of HA and LA children had to be observed (and tested) within a limited period of time, and because it seemed apparent to us that the development of a refined methodology would not be a task of short duration, we decided to observe each child for one hour at the end of which time the observer was (a) to describe the major events which had occurred and (b) to write as integrated a personality description as was possible from such a brief observation.

SUBJECTS

The 32 pairs of HA and LA children were matched for sex, grade, and locally standardized T-scores on the Otis tests of mental ability. They differed in that the HA subjects were drawn from the top quartile of scores

on both the test anxiety and general anxiety scales while the LA subjects were drawn from the bottom quartile. The scales were administered in the previous year when the children were in grades 1 to 4; the observations were made when the children were in grades 2 to 5. In each anxiety group there was an equal number of boys and girls. The 32 matched pairs were the same as those used in the Rorschach study (2). In the learning study (5), and for reasons given there, it was only possible to use 24 of the matched cases.

PROCEDURE

Two observers were used, and each observed half of the sample. For purposes of establishing observer reliability, 16 of the subjects were observed independently by each examiner on successive days at the same hour of day. As was indicated earlier, each child was observed for approximately one hour after which the description and interpretation of behavior was written. After all children were observed, and the reliability of the observations evaluated, an abstract of each description was prepared. This abstract contained words, phrases, or events which either occurred with greatest frequency in the report or which clearly represented the observer's dominant reaction to the child. The data to be presented below (with the exception of that concerned with reliability) are based on frequency counts of words, phrases or events contained in these abstracts. The observers were never aware of the group (HA or LA) to which a subject belonged.

The HA and LA groups were compared on such variables as need achievement, unintelligible behavior, academic ability, task orientation, motor behavior, and teacher relationships. It should be noted that the observations (and the abstract prepared from them) did not necessarily contain the above phrases but they have here been used because they clearly reflected the words and phrases actually used by the observer. In addition, and as was indicated previously, the presence of these words or phrases in the observations reflected the observer's opinion that these were strong or distinctive characteristics of the child. For example, there may have been times during an observation when a child's behavior was seen by the observer as either "dreamy," or "staring," or "baffling," or "puzzling," but unless the observer had stated in her report that this was a dominant characteristic of the child it was not contained in the abstract and, consequently, he was not counted among those with "unintelligible behavior."

Statistical tests have been omitted in presenting our results, primarily because the number of cases involved in many of the comparisons was, in our opinion, too small to warrant such procedures. If the observations had been based on an explicit schedule so that each of our subjects could have been scored or rated on every variable, the problem of the interpretation of results based on small numbers in each cell of an analysis would

have been less difficult. For reasons given earlier the observations were not conducted in this manner.

RESULTS

Reliability

In the reports of the 16 cases used for establishing reliability, identifying characteristics were deleted. These cases were divided into groups of four, and for each group there were eight reports, two for each child. The groupings were done to minimize differences in intelligence and grade within each group. Within each of the four groups each observer had to match her report with that of the same child made by the other observer. One observer correctly matched each of her 16 reports with those of the other observer. The other observer was successful in 14 of the 16 cases. These results were far better than was expected. From observation of the matching behavior of the observers, as well as discussions with them after the matchings were completed, there seemed no basis for assuming that their degree of success could be attributed to what may be termed artifacts in the reports. A third judge was then given the protocols and he successfully matched 11 of the cases. A fourth judge, who was only able to complete two groups of matchings (i.e., eight cases), matched six of the eight cases correctly. We conclude from these results that the observations of the two observers had satisfactory reliability. It should be pointed out, however, that in most, if not all, of the incorrect matches made by the three judges (one judge had no incorrect matchings) the child's behavior in the two observation periods seemed markedly different.

Need Achievement

There were eight LA and seven HA boys who can be described as having a pervasive need for achievement either because they were strikingly attentive to tasks, or because of active, eager, energetic and frequent volunteering, or carefulness of work, or evidence of pleasure in good work, or concern with lessons. More often than not the observations of these boys contained more than one of these phrases. Whereas in the case of boys there was obviously no difference between anxiety groups, in the case of girls there were eleven HA and three LA whose observed behavior suggests a strong need for achievement. There were no HA girls who were explicitly described as "not particularly interested or motivated" while six LA girls were so described; among the boys three LA and four HA were so described. It would seem, therefore, that observational data suggestive of a strong achievement need tended to distinguish between HA and LA girls but not between HA and LA boys.

"Unintelligible" Behavior

In stating that a child has a strong need to achieve, one is clearly relating his behavior to some external task or situation—at least, one

assumes that one understands why a child is responding in a particular way. There were more than a few children whose behavior was unfathomable by the observer and for which such adjectives as "dreamy," "staring," "baffling," "puzzling," and "enigmatic" were used. Among the boys three HA and four LA were so described; among the girls eight LA and four HA were so described. The fact that much of the behavior of 50 per cent of the LA girls could not be interpreted by the observer is undoubtedly related to the previously described differences in achievement behavior between HA and LA girls.

References to Academic Ability

Of the 16 LA boys eight were explicitly labeled as either superior or adequate in academic ability, one as somewhat inadequate; in seven cases there was no mention at all of academic ability. Of the HA boys three were labeled as either superior or adequate, five as having an academic problem; there was no reference to academic ability in eight cases. There was no trend at all in the distribution of these observational labels between HA and LA girls. These results suggest that in contrast to the LA boys the academic ability or performance of the HA boys was less impressive, whereas there was no difference between HA and LA girls. Not only was no LA boy described as an academic problem, but more LA than HA boys behaved in a way so as to cause the observer to record either their presumed adequacy or superiority.

Task Orientation

This variable refers to the degree to which the child could maintain attention to the different tasks presented him. The behavior relevant to this variable was categorized as follows:

1. Task oriented. The child's attention very rarely, if at all, was distracted from tasks.
2. Slight variability. Similar to "task oriented" but child was sometimes distracted from a task, or allowed himself to be distracted. The amount of attention given to different tasks varied slightly (e.g., between reading and arithmetic).
3. Marked variability. The child was very attentive sometimes but very distracted at other times.
4. Generally distractible. Attention to classwork was negligible.

It seems clear from Table 1 that the HA boys again presented a less favorable picture (or created a less favorable impression on an observer) than did LA boys. If one were to combine the "task oriented" and "slight variability" categories, there would still be a tendency (fourteen LA vs. nine HA) for the LA boys to give a picture of greater efficiency in performance. The major difference between HA and LA girls is that more of the latter were found in the "generally distractible" category. If one were

CHILD DEVELOPMENT

TABLE I

DIFFERENCES IN TASK ORIENTATION BETWEEN HA AND LA GROUPS

	<i>Task Oriented</i>	<i>Slight Variability</i>	<i>Marked Variability</i>	<i>Generally Distractible</i>
LA boys ..	10	4	0	2
HA boys ..	4	5	5	2
LA girls ..	5	0	5	6
HA girls ..	4	5	5	2

to combine the first two and last two categories, one finds again that the more favorable impression was created by the HA girls.

Motor Behavior

The original observations of each child were scrutinized and a tally made of various types of motor behavior (e.g., nail biting, tongue movement, mouthing noises, wiggling, etc.). There was a marked tendency for HA boys, in contrast to LA ones, to manifest more different types of motor behavior, whereas the reverse tendency was the case between HA and LA girls. Unfortunately, however, it turned out that one of the two observers noted motor behavior in her observations to a much greater extent than the other—an observer difference which did not occur with other variables we have discussed. Although the results of the motor behavior analysis are compatible with our other findings, particularly in the case of the HA and LA boys, we are clearly not justified in doing anything other than mention the findings and their ambiguity in light of the observers' different degrees of attention to motor behavior.

Teacher Relationships

Characterizing the child's attitude toward a relationship with the teacher is difficult because of the different ways in which such attitudes or relationships were manifested. During the course of a single observation the child might be seen as having different kinds of relationships with the teacher. For example, at one point a child may ask the teacher an unusual number of questions while at another point he may be misbehaving in order to get attention—both types of behavior being phenotypically different while perhaps being genotypically the same. For the purposes of the present study, however, we have put a child into what was for him the most frequent category. Among the HA boys there were six cases who tended to ask an unusual number of questions of the teacher, suggestive either of dependency or attention-getting. There were three such cases among the LA boys. There were four HA boys whose behavior changed markedly when the teacher was near them in that they would then make a display of concentrated study—or in some way acted as if hoping to

avoid the teacher's attention; in three of these cases the description seemed to suggest that the child feared the possibility of censure by the teacher (one of these cases was explicitly described as "scared"). There was no such case among the LA boys. There were, therefore, ten HA and three LA boys whose overt behavior in relation to the teacher suggested insecurity. In the remaining cases in each of these two groups the behavior in relation to the teacher was ambiguous or, more frequently, the observations simply did not contain material relevant to the problem. Among the girls the content of the observations was of a different quality. For example, whereas nine boys characteristically asked many questions of the teacher, only one girl was so described. The kinds of differences observed between HA and LA boys were not observed between HA and LA girls.

DISCUSSION

The major question arising from our results is why expected differences were obtained between HA and LA boys and not between HA and LA girls? In contrast to LA boys with whom they were matched for grade and IQ, the HA boys did not seem as academically impressive, their attention to tasks not as persistent, and their relationship to the teacher more suggestive of feelings of personal and academic inadequacy. In the Rorschach study (2), in which these same children participated, the HA boys, in comparison to LA ones, were also behaviorally more dependent and insecure—differences which were not found between HA and LA girls. In the experimental learning study (5) with these subjects the HA boys performed significantly more poorly than their LA counterparts, whereas no significant differences were found between HA and LA girls, although the trend was in the same direction.

There is one respect in which the results of the classroom observations differ from previous findings. Whereas we have previously found no marked differences between HA and LA girls, in the classroom observations it was the HA girl in contrast to the LA one who manifested a greater need for achievement, less "unintelligible" behavior, and more persistent attention to tasks at hand. Behaviorally the LA girls seemed more similar to the HA rather than the LA boys. Why, therefore, does not one get differences between HA and LA girls on performance criteria? It is conceivable that HA girls are not as anxious as their anxiety scores suggest and/or that LA girls are not as free from interfering modes of response as their scores would suggest.

A possible explanation of our various findings with our 32 matched pairs is a sampling one: we could have been, so to speak, less fortunate in choosing our girls than our boy subjects. We had anticipated the need for replication and had planned to carry out in England, in connection with our studies there (4), studies identical to those done in our country. If similar results could be obtained cross-culturally, it would reduce the

cogency of an explanation based on sampling. This turned out not to be possible and, instead, replications in the United States are planned. However, other studies (2, 4, 5) which we have done with our anxiety scales strongly suggest that our findings with our matched pairs may involve factors other than sampling ones. Our present thinking about the problem involves several hypotheses:

1. In our culture, at least, it is generally more difficult for boys than for girls to admit to anxieties, worries, or fears. While we expect and support such admissions in girls, we do not so react to similar behavior in boys. It is these hypotheses which we think helps explain why girls consistently get higher scores on our anxiety scales.

2. High anxiety scores are relatively uncommon among boys and when such scores are obtained they reflect negative self-attitudes and modes of responding which might be termed interfering. One would expect, therefore, that HA and LA boys would be differentiated from each other on performance criteria.

3. High anxiety scores are not uncommon among girls and when such scores are obtained they do not necessarily reflect negative self-attitudes or interfering modes of response. In addition, and as the classroom observations suggest, high scores among girls may reflect an unusually high level of motivation to achieve which does not have adverse effects on performance. One would not necessarily expect, therefore, that HA and LA girls would be differentiated from each other on performance criteria.

4. Low anxiety scores are somewhat unusual among girls (certainly more unusual among girls than boys) and it may be assumed that some of the low-scoring girls have an unusually strong need against admitting anxiety, reflecting, perhaps, a stronger anxiety than in the case of a low-scoring boy whom we expect to "hide" his anxiety. This would be an added factor working against differentiation between HA and LA girls.

5. Finally, it may be that the content of our scales taps areas of anxiety more pertinent to boys than to girls. It is our impression, based primarily on interviews with parents, that eliciting and maintaining positive affective reactions from significant adults is more important to girls than to boys, and thus more likely to be accompanied by anxiety in girls. If, for example, one could devise an anxiety questionnaire concerning attitudes toward and experiences with teachers and other significant adults, one might then find significant differences in performance between HA and LA girls.

Studies are being planned to evaluate these hypotheses. However, it is our present opinion that concentration on the last hypothesis will prove to be the most fruitful.

SUMMARY

This study involved the direct classroom observations of high (HA) and low anxious (LA) children, anxiety being defined by scores on test

anxiety and general anxiety scales. There were 32 pairs of subjects matched for grade, sex, and IQ score but differing in anxiety score. Satisfactory reliability of observations was established. HA boys differed from LA boys in that they did not appear as academically adequate, showed less task orientation, and in relationships with the teacher manifested greater insecurity. HA girls differed from LA girls in that they seemed to have a stronger need for achievement, showed less "unintelligible" behavior, and contained fewer distractible individuals. The relationships of these findings to those obtained on the same subjects in other studies are discussed. Several hypotheses were advanced to account for the differences in the correlates of anxiety scores in boys and girls.

REFERENCES

1. DAVIDSON, K. S., SARASON, S. B., LIDTHALL, F. F., WAITE, R. R., & SARNOFF, I. Differences between mothers' and fathers' ratings of low anxious and high anxious children. *Child Develpm.*, 1958, 29, 155-160.
2. SARASON, S. B., DAVIDSON, K. S., LIDTHALL, F. F., & WAITE, R. R. Rorschach behavior and performance of high and low anxious children. *Child Develpm.*, 1958, 29, 277-285.
3. SARASON, S. B., DAVIDSON, K. S., LIDTHALL, F. F., & WAITE, R. R. A test anxiety scale for children. *Child Develpm.*, 1958, 29, 105-113.
4. SARNOFF, I., LIDTHALL, F. F., WAITE, R. R., DAVIDSON, K. S., & SARASON, S. B. A cross-cultural study of anxiety amongst American and English school children. *J. educ. Psychol.*, in press.
5. WAITE, R. R., SARASON, S. B., LIDTHALL, F. F., & DAVIDSON, K. S. A study of anxiety and learning in children. *J. abnorm. soc. Psychol.*, in press.

HUMAN FIGURE DRAWINGS OF HIGH AND LOW ANXIOUS CHILDREN¹

CYNTHIA FOX, KENNETH DAVIDSON, FREDERICK LIGHTHALL,
RICHARD WAITE, and SEYMOUR B. SARASON

Yale University

This study explores the differences between high anxious (HA) and low anxious (LA) children in their drawings of human figures. It is part of an ongoing project investigating the measurement and correlates of anxiety in school children. The construction and validation of the anxiety measures, the test anxiety scale and the general anxiety scale, have been described in earlier reports from the project (1, 3, 4).

The testing procedure involved administration of the two scales in a single group session to a large sample of subjects. Interpolated between the two scales, to reduce possible fatigue and boredom, was the task of drawing a boy, drawing a girl, and drawing a house. It was our expectation that some reflection of a child's level of anxiety or of correlates of his anxiety level would appear in this sort of independent, semistructured, expressive task. In this study, only the human figure drawings were considered.

METHOD

Subjects

Subjects consisted of three distinct samples drawn from 747 first through fourth grade children who had taken the anxiety questionnaires. Each of the first two samples, which were used for exploratory work, consisted of those boys and girls from two randomly selected schools who were in approximately the upper and lower quartiles on the general anxiety scales. Each of these samples contained about 40 subjects. The third and final sample, for which results will be reported, consisted of 32 pairs of subjects matched for grade, sex and average verbal and nonverbal T-score

¹ The project of which this study was a part is being supported by a grant from U.S.P.H.S.

on the Otis Quick Scoring Mental Ability Test (Alpha). One member of each pair of subjects was in the fourth quartile of scores on Test Anxiety and General Anxiety scales while the other was in the first quartile. There was an equal number of boys and girls in each anxiety group. Overlap between the three samples was avoided.

Procedure

The drawings in the first sample were examined with no knowledge of the anxiety scores of the subjects and an attempt was made to judge each child's set of drawings as HA or LA. Predictive accuracy was very poor. Following this, a list of about 20 variables of the drawings was drawn up which seemed likely to correlate with anxiety level. These variables were selected on the basis of theoretical expectations, reports in the literature, and careful examination of the drawings in sample I. Each set of drawings in sample I was scored on a two-point scale for presence of each variable. Of the initial set of variables, nine seemed promising in discriminating between HA and LA subjects. The drawings in sample II were scored with respect to each of these nine variables and six discriminated significantly between HA and LA subjects. It was only at this point that sample III, consisting of the 32 matched pairs of children, was examined. The experimenter first made the judgment of HA or LA for each set of drawings on the basis of total clinical impression. With no knowledge of the accuracy of these predictions, she scored the drawings for each of the six variables that emerged from work with samples I and II.

Reliability of scoring was established by having another investigator, who had not previously worked with the drawings, score the drawings in sample II for the six discriminating variables. After sample II had been scored, the two scorers discussed disagreements and tried to clarify scoring criteria. Sample III was then scored by the second investigator. It should be noted that, with the exception of work with sample I, all judgments were made without knowing whether subjects were HA or LA.

RESULTS AND DISCUSSION

The results reported here derive from work with sample III. The six variables which were scored and their definitions are as follows:

1. *Mutilation*—scored as present if one or more limbs or facial features (eyes, nose, mouth) is absent; or if either ears, hands or feet are absent when they had been included in one of the two drawings. Thus, mutilation was scored if a child had drawn hands on one figure but not on the other. It was not scored if no hands were drawn at all. Mutilation was also scored if one or more limbs was *markedly* small compared to the rest of the body.
2. *Smile*—scored present if the corners of the mouths in both drawings turned upward or if the corners of the mouth in one of the drawings

turned up and the other was ambiguous. Scored absent if both were ambiguous or if one or both turned down.

3. *Shading*—scored present if there was any blackening in of portions of the drawings. The one exception was hair which was filled in by most of the children and therefore not scored.

4. *Arm position—down*—scored present if one or more arms made less than a 45 degree angle with the body or turned in toward the body.

5. *Rigidity*—scored present if the figures appeared rigid, unable to move, likely to topple over if they did move.

6. *Playfulness—humor*—(most subjective of the judgments) scored present if a particular detail or some expressive stance of the figure communicated a kind of playful, humorous mood. (This was not the antithesis of rigidity since drawings scored as playful included some that were also scored rigid.)

All of the variables discriminated significantly between HA and LA subjects in the directions predicted on the basis of results from samples I and II. Thus, HA subjects tended to have more mutilation and more rigidity in their drawings than LA subjects. Smiling, arm-position-down, and playfulness-humor tended to occur more often in the drawings of LA subjects than in the drawings of HA subjects. Shading was more complex. As was predicted, shading tended to be associated with high anxiety in boys and low anxiety in girls. A chi square analysis was used to test the significance of the differences and results are presented in Table 1. Since the directions of the relationships were predicted, a one-tail probability value is reported. Scoring reliability is reported in terms of the phi coefficients and is also presented in Table 1.

TABLE 1

RELIABILITY OF SCORING, AND DIFFERENCES BETWEEN HA AND LA CHILDREN IN FIGURE DRAWING VARIABLES

Variable	Directional Difference	p	Scoring Reliability
Mutilation	HA > LA	<.02	.84
Smile	LA > HA	<.01	.86
Shading	HA boys > LA boys	<.02	.92
	LA girls > HA girls		
Arm position—down ..	LA > HA	<.02	.89
Rigidity	HA > LA	<.01	.79
Playfulness—humor ...	LA > HA	<.01	.77

The attempt to judge anxiety level from an over-all examination of the drawings resulted in 46 correct placements ($p = .02$). It should be noted that it was the judge's impression that her assignments were based on an over-all impression, the determinants of which are difficult to specify. While there undoubtedly was some consideration of the six variables in making these judgments, they did not seem primary. When considered in the light of the judge's inability to discriminate HA and LA drawings at the beginning of the study (Sample I), this finding suggests that significant learning did occur as a result of familiarization with this type of material. Thus, having looked at enough drawings and compared them with anxiety scores, the judge could learn to discriminate between the productions of subjects who scored HA and subjects who scored LA.

Explanations of the relation between the six variables and anxiety level must remain speculative in an empirical correlational study of this sort. At a speculative level, however, it seems that the three variables which correlated negatively with anxiety (presence of smile, arm position-down, and humor-playfulness) reflect a degree of freedom to become involved in a creative task in an expressive, pleasurable, and nonconstricted fashion. In contrast, the tendency for these features to be absent in the drawings of HA children along with the tendency for mutilation and rigidity to be present suggests that in the HA child, accuracy, spontaneity, and expressiveness are interfered with, at least in a semistructured creative situation of this sort, and that for him it is a more threatening and less pleasurable experience than for the LA child. The increased presence of mutilation in the drawings of HA children suggests, in addition, some concern with and questioning of body integrity and adequacy. It is consistent with an earlier finding of more anatomy responses in the Rorschach records of HA children (2). We have no ready explanation for the different direction of relationship between shading and level of anxiety in boys and in girls. This finding is consistent, however, with earlier results (1, 2, 4, 5) which suggest that some of the correlates and modes of expression of anxiety may be different for boys than they are for girls.

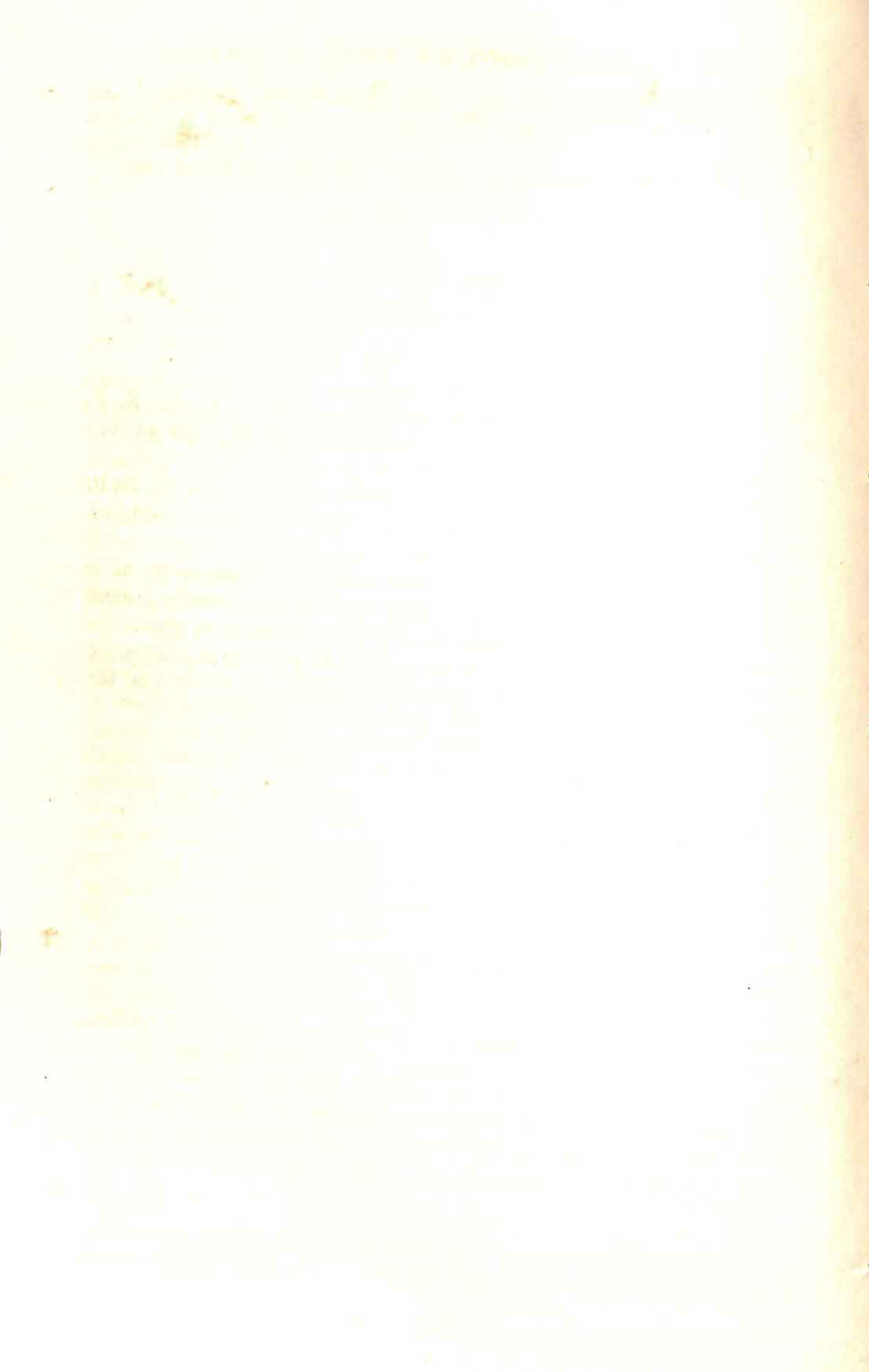
SUMMARY

This study concerned the human figure drawings of high (HA) and low (LA) anxious children, anxiety being defined by scores on test anxiety and general anxiety scales. Thirty-two pairs of subjects were used, matched for grade, sex and IQ score but differing in anxiety scores. The variables scored were determined by analyses of drawings of two independent "pretest" samples. Scoring reliability was satisfactory. The six variables scored with the final sample of 32 pairs differentiated between HA and LA subjects as they did in the two pretest samples. The drawings of HA subjects showed significantly more "mutilation" and "rigidity" than did those of LA subjects. The drawings of LA subjects, in contrast to those of HA

subjects, showed significantly more "playfulness-humor," "smiling," and "arm position down." HA boys had more shading than LA boys while LA girls had more than HA girls, again as found in the two pretest samples.

REFERENCES

1. DAVIDSON, K. S., SARASON, S. B., LIGHTHALL, F. F., WAITE, R. R., & SARNOFF, I. Differences between mothers' and fathers' ratings of low anxious and high anxious children. *Child Developm.*, 1958, 29, 155-160.
2. SARASON, S. B., DAVIDSON, K. S., LIGHTHALL, F. F., & WAITE, R. R. Rorschach behavior and performance of high and low anxious children. *Child Developm.*, 1958, 29, 277-285.
3. SARASON, S. B., DAVIDSON, K. S., LIGHTHALL, F. F., & WAITE, R. R. A test anxiety scale for children. *Child Developm.*, 1958, 29, 105-113.
4. SARNOFF, I., LIGHTHALL, F. F., WAITE, R. R., DAVIDSON, K. S., & SARASON, S. B. A cross-cultural study of anxiety amongst American and English school children. *J. educ. Psychol.*, in press.
5. WAITE, R. R., SARASON, S. B., LIGHTHALL, F. F., & DAVIDSON, K. S. A study of anxiety and learning in children. *J. abnorm. soc. Psychol.*, in press.



THE NEED FOR SIMPLICITY IN RESEARCH IN CHILD PSYCHOLOGY¹

GLENN TERRELL

University of Colorado

Perhaps the best way to begin this paper is to explain what the writer means by simple research in child psychology. A few examples should suffice: (a) the effects of interference on crank turning behavior (16); (b) an experiment involving the learning of concepts of size (6); (c) an investigation of reinforcement variables affecting the resistance to extinction of a simple response of placing a ball in a hole (1). In each of these problems the behavior studied is relatively simple, particularly in relation to far more complex responses involved in studies of social development, or in investigations of the effects of child rearing methods on personality growth. It is felt that the best way to begin the argument for the current need for simple research in child psychology is to make an analysis of a characteristic area of research involving exceedingly complex variables. For this purpose, the writer has selected the general area involving the effects of child rearing variables on personality development. Out of this analysis and subsequent argument should come the impression that the writer's main reason for favoring "basic" research is a methodological one. Following this discussion, the writer will present some of the important uses of the study of simple processes in child psychology, and follow this with a treatment of the advantages that such a research orientation has.

The research literature in child development is replete with investigations concerned with the relationships between early guidance or child rearing procedures and personality development of children. The independent and dependent variables of these studies are typically complex, hard-to-measure processes. Many of them involve attempts to measure the effects of such variables as punishment, affection, rejection, breast feeding, and home atmosphere on the adjustment of the child. Measures of these child rearing variables are usually obtained through interviews, rating scales, projective tests, or observational procedures. The adjustment or response measures are typically scores on paper and pencil personality tests, clinical judgments, teacher ratings, or responses in projective doll play situations.

¹ This paper is a longer version of a contribution to a symposium, "Trends in Developmental Psychology," presented at the Rocky Mountain Psychological Association meetings in 1957.

Correlations between these child rearing and adjustment measures are calculated, and if they differ significantly from zero, the investigator typically concludes that he has discovered a useful relationship in the science of child rearing.

Much has been written about the unreliabilities of the interview, the rating scale, personality tests, and clinical judgments. It is not within the scope of this paper to go into a detailed review of the difficulties associated with the use of these methods. Suffice it to say that there are problems of recall of the desired information by the parent, of intentional or unintentional falsification of responses, of definition, and of projections or misinterpretations by clinicians, parents, and teachers.

To illustrate what the writer considers the prematurity of research at this complex level in child psychology, perhaps the analysis of an hypothetical example will prove helpful. Let us suppose that an investigator wishes to study, among other things, the relationship between the duration of breast feeding in infancy and the frequency of dependency behavior of elementary school children. Measures of duration of breast feeding are obtained by the interview, while measures of dependency are taken in a standardized projective doll play situation. For the sake of the argument, let us assume that the investigation is conducted with the greatest possible care. Much preparation has gone into the interview and the doll play procedures in order that the researcher reduce to a minimum the invalidities and unreliabilities of the study due to these factors. Further, let it be assumed that the investigator finds a significant inverse correlation between duration of breast feeding and the frequency of dependency behavior.

To what extent is he justified in claiming the discovery of a relationship between these variables? Obviously the experimenter is not manipulating the independent variable. Though a relevant variable can be isolated and studied without the experimenter's manipulating it, the difficulty with this procedure, as McCandless and Spiker (11) recently pointed out in an appeal for more experimental research in child psychology, is greatly increased as the number of relevant variables affecting the dependent system is increased. And as McCandless and Spiker in the same paper further point out, "That a very large number of variables affects the behavior of the living organism is a fact that has long been recognized" (p. 76). Here is the crux of the difficulty: the number of relevant variables affecting the dependent system in complex research, as in the hypothetical example, is undoubtedly far greater than those affecting relatively simple responses. Furthermore, they are infinitely more difficult to isolate and control. In fact, the control of variables associated with the study of the effects of child rearing techniques on adjustment is in most cases impossible at the present stage. Indeed, it is highly probable that the great majority of these variables is as yet unidentified. Some of the possible variables more obviously relevant to dependency, other than the duration of breast feeding, are the character structure of the parents, friendship patterns, socioeconomic class,

genetic and constitutional factors, and the nature of sibling relationships. Variables such as these are more often than not unconsidered in studies at this level, and even if efforts to study them were made, the outlook would indeed be bleak, since it would be difficult if not impossible to say which has affected the dependent variable, and to what degree. The end result of this state of affairs is, of course, that the experimenter thinks he has discovered a relationship in child rearing variables, when in reality there is a good chance that one or more of the unconsidered variables mentioned above has brought about the change in the dependent variables. Any well trained scientist must admit that because of the above reasoning, there is, at the very best, only a slight, highly equivocal suggestion in the results of the above described hypothetical study that there is a meaningful relationship between the variables studied, certainly not enough to warrant the confident statements about child rearing problems one frequently sees emanating from such research. It is not at all unlikely that many child psychologists have been misled into conducting further research along these lines, to say nothing of the possibly more serious consequences of parents being misled by reports of these studies.

At this point it seems appropriate to point out some important uses of more basic research at the present stage of maturity of developmental psychology. If one has followed recent issues of *Child Development*, *The Journal of Experimental Psychology*, and other psychological journals, it is unmistakably clear that one of the newer trends in developmental psychology is a preoccupation with basic research problems that relate to some of the significant problems in general psychology, particularly within a stimulus-response framework. Two of these problems will be described briefly, along with some of the research that has been performed with children bearing on these issues.

First, there is the principle of stimulus generalization, which has been used in general behavior theory to explain, among others, such phenomena as the occurrence of sudden or insightful solutions of problems and the persistence of behavior that is not rewarded, or is even punished. One of the important variables assumed to affect the generalization of a response is the number of reinforcements given on the training stimulus. Razran (14) reports that the evidence from classical conditioning with infrahuman and adult human subjects demonstrates fairly conclusively that an increase in the number of reinforcements on the training (conditioned) stimulus results in an increased amount of generalization to the test stimuli. Spiker (21) has developed a simple technique to extend these findings to the young child. The Ss, preschoolers, were trained to pull a lever repeatedly to a stimulus of a given hue in order to receive marbles. They were then tested on stimuli of different hues to determine the number of responses that would be made to the test stimuli without receiving reinforcement. A group given 24 reinforcements to the conditioned stimulus prior to the generalization test averaged nearly twice as many responses to the test

stimuli as did a group given only 12 preliminary reinforcements. Thus, the same relationship found to exist at the infrahuman and human adult levels was found to apply to the preschool child. Spiker (22) in a later experiment found that the steepness of the generalization gradient was also a positive function of the number of reinforced training trials as well as the intensity of the stimuli used in the training. In this experiment 60 Ss were given differential reinforcement to a white (positive) and a blue (negative) light. For one half the Ss, the positive stimulus was the brightest of four stimuli differing from each other in brightness; for the other half, it was the dimmest. One half of each of these groups was given 12 presentations and the other half was given 24 presentations of each the positive and negative stimuli. The response was again the repeated pulling of a lever for marbles during a 3-second presentation of the conditioned stimulus, and the response measure was the number of such responses that occurred during the 3-second periods. Immediately following training, the Ss were tested, without reinforcement, on each of the four stimuli differing in brightness. As was mentioned above, the findings were: (a) a steeper gradient for the Ss who received the greater number of reinforced training trials, and (b) a steeper gradient for the bright-trained than for the dim-trained Ss. These results are in agreement with the experiments by Grice and Saltz (5) and by Brown (2) at the infrahuman level, and with the prediction made by Hull (8) on the basis of the principle of stimulus dynamism.

The second principle from general psychology which has received recent attention in child psychology is the principle of secondary or mediated generalization, sometimes referred to as acquired equivalence of cues. In brief, this principle states that if S has been trained to make the same response to two or more dissimilar stimuli, there will be an increased tendency to generalize to the other stimulus other responses subsequently learned to one of the stimuli. In other words, if S is taught to make response 1 to stimuli 1 and 2, then taught to make response 2 to stimulus 1, he will make response 2 to stimulus 2. Jeffrey (9) has demonstrated the role of mediating responses in generalization. He trained his Ss to move a lever, again a simple dependent variable, in one direction to a white stimulus, and in the opposite direction to a black stimulus. He then taught some Ss to call a gray stimulus "white" and others to call it "black," and next retrained the Ss on the lever moving task to the white and black stimuli. Following this, he presented the gray stimulus interspersed with black and white stimuli to determine how the Ss would respond to the gray stimulus. He found that, if the Ss had been taught to call the gray stimulus "white," they responded to it as they did to the white stimulus; if they had been taught to call the gray stimulus "black," they responded to it as they did to the black stimulus. He found a similar tendency for another group of Ss who instead of using the names "black" and "white" had been taught to turn the handle to the right or the left for the white, black and gray stimuli. Thus, Jeffrey demonstrates that either verbal or

motor responses may serve as mediating responses. Shepard (18) and Eisman (3), among others, have also recently demonstrated mediated generalization in young children.

The study of simple responses like the ones described in the experiments above may be made in more applied, pragmatic research with children. For example, in the important area of incentives in children's learning, the writer and Kennedy (24) have recently shown that a candy reward results in quicker learning and more consistent transfer of a "larger-than" concept than either praise, reproof, a light flash, or a delayed reward. The only exception to this latter statement was the nonsignificant difference that existed in the consistency of transfer between the candy reward group and the group that was given a delayed reward. In this experiment the behavior studied was a simple button-pushing response to the larger of two three-dimensional geometric objects. The nonsignificant difference in the candy (immediate) and delayed reward group of the aforementioned study led the writer (23) to perform an experiment comparing two types of delayed rewards with each other and with a immediate reward in the learning and transferring of a "larger-than" discrimination. It was hypothesized that Ss assigned to a delayed reward condition which permitted them to observe progress toward the to-be-received reward would learn more quickly and transfer more consistently than would Ss assigned to a delayed reward condition that did not allow them to observe their progress in the learning situation. This hypothesis was convincingly supported by the results of the experiment.

Now what advantages do the studies of simple behaviors in the experiments cited above have, other than their contributions to general psychology? As was stated previously, they are mainly methodological. First of all, when we study basic responses, the variables, both independent and dependent, are easier to manipulate. Compare, for example, the difficulties associated with isolating, controlling, and measuring complex child rearing and adjustment variables with the relative ease and rigor of control in (a) Jeffrey's study of the effects of simple verbal instructions, "say white" or "say black," on crank turning responses, or (b) Spiker's experiment with the generalization of lever pulling responses to lights of varying hues, intensities, and frequencies. Because of the more rigorous control exercised in the latter experiments, the experimenter minimizes the danger of masking the basic processes. That we can be less equivocal about interpreting the results of these experiments is obvious.

In most experiments the researcher has a choice in the matter of selecting responses to be studied. For example, Spiker may well have studied the principle of stimulus generalization within a more complex, clinical framework, by obtaining measures of generalization of responses of children from maladjusted teachers to peers or love interests, instead of studying the generalization of lever pulling responses to lights of differing intensities and hues. Or again, the writer may have studied the effects of various

incentives on the solution of higher mathematics problems, or on some of the more intricate aspects of social development in childhood. That we are all interested in the latter problems is a foregone conclusion. In fact, the critics of the writer's position will ask the question, "Of what practical value is it to investigate these simple, trivial behaviors? What we need to know, indeed what society is demanding of us now, is the solution of important practical problems in developmental psychology, such as the complex factors involved in learning and personality development." The writer's answer to this argument is that he believes that ultimately we will be able to supply answers to the complex, immediately relevant social problems associated with child psychology more quickly if we first systematize our knowledge of relationships existing at an elementary behavior level. By doing this we do not get ahead of our limited methodological resources, which has always been one of the greatest tendencies of researchers in child psychology. And what is even more important, we do not run the serious risk of misleading society.

Critics of this position will insist that the writer be more specific about *how* the study of relatively uncomplicated problems eventually will facilitate a rigorous, more fruitful investigation of such phenomena as are involved in complex learning or personality development. It is felt that not only will the current concern for basic research with children mean that we keep our investigations within the limits of our methodological resources, but it is also believed that, because of the above, basic research is likely to result in the discovery of variables relevant to complex human behavior. For example, the writer (23) has found a strong indication that the speed with which a child learns a simple "larger-than" discrimination without a material reward depends upon a host of child rearing variables associated with social class membership, such as the importance to the middle class child of not appearing unintelligent. Had the relevance of this social class variable been suggested by an experiment involving more complex learning with less rigor of control, it appears reasonable to argue that one would place less confidence in its importance, even in complex learning situations. It is true that a variable related to the learning of a size discrimination may not be important in the learning of principles involved in the solution of quadratic equations, or the variables relevant to the generalization of responses from lights of one intensity to lights of other intensities may not be related to the generalization of responses from maladjusted parents to heterosexual relationships. There is some evidence, however, for contending that the likelihood of variables which are known to apply at a simple behavioral level to be relevant also at a more complex level is greater than the probability that variables suggested as relevant in complicated, poorly controlled studies are in fact relevant at that level. The evidence in support of this statement falls into two categories: (a) the applicability of principles discovered in animal research to more complex human behavior, and (b) the highly inconsistent findings of research in human

behavior concerned with such complex phenomena as child rearing variables and their effects on personality development.

Rosenblum (15) has found that the acquisition and extinction of "copying" responses by fourth and fifth graders are regulated by the same principles of partial reinforcement as found in animal experiments (20). Kuenne (10) reports that Spence's (19) theory of transposition, formulated to account for animal behavior in a discrimination learning situation, also applies to the preverbal child. Consider also the highly controversial research involving the effects of socialization training on personality. Some studies show the techniques of feeding, weaning, and toileting to be related in a substantial way to security, the quality of adjustment, and other important personality traits (4, 7, 12); whereas others show no relationship between these variables (13, 17, 25). It is felt that the inconsistencies in these investigations are due, among other things, to false leads concerning relevant variables in socialization and personality that researchers get in focusing attention on problems of this degree of complexity.

The writer does not want to take the position that *all* research in complex human behavior is useless at the present stage of maturity of psychology. It must be acknowledged that some of this research is yielding reasonably consistent results, despite the difficulties discussed above. It is felt, however, that child psychology generally would benefit more by an increased concern on the part of many of us with understanding behavior in its more basic forms. Indeed, it is felt that the great preponderance of the research which is now being done at a more complicated level will have to be redone in the light of the results of the more basic research in child psychology.

REFERENCES

1. BIJOU, S. W. Patterns of reinforcement and resistance to extinction in young children. *Child Developm.*, 1957, 28, 47-54.
2. BROWN, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, 33, 209-226.
3. EISMAN, B. S. Attitude formation: the development of a color preference response through mediated generalization. *J. abnorm. soc. Psychol.*, 1955, 50, 321-326.
4. GOLDMAN-EISLER, F. Breast feeding and character formation. In C. Kluckhohn & H. A. Murray (Eds.), *Personality in nature, society and culture*. New York: Knopf, 1953, Pp. 146-184.
5. GRICE, R. G., & SALTZ, E. The generalization of an instrumental response to stimuli varying in the size dimension. *J. exp. Psychol.*, 1950, 40, 702-708.
6. HICKS, J. A., & STEWART, F. D. The learning of abstract concepts of size. *Child Developm.*, 1930, 1, 195-203.

7. HOLWAY, A. R. Early self-regulation of infants and later behavior in play interviews. *Amer. J. Orthopsychiat.*, 1949, 19, 612-622.
8. HULL, C. L. *A behavior system*. New Haven: Yale Univer. Press, 1952.
9. JEFFREY, W. E. The effects of verbal and non-verbal responses in mediating an instrumental act. *J. exp. Psychol.*, 1953, 45, 327-333.
10. KUENNE, M. R. Experimental investigation of transposition behavior in young children. *J. exp. Psychol.*, 1946, 36, 471-490.
11. McCANDLESS, B. R., & SPIKER, C. C. Experimental research in child psychology. *Child Developm.*, 1956, 27, 75-80.
12. MASLOW, A. H., & SZILAGYI-KESSLER, I. Security and breast feeding. *J. abnorm. soc. Psychol.*, 1946, 41, 83-85.
13. PETERSON, C. H., & SPANO, F. Breast feeding, maternal rejection, and child personality. *J. Pers.*, 1941, 10, 62-66.
14. RAZRAN, G. Stimulus generalization of conditional responses. *Psychol. Bull.*, 1949, 46, 337-365.
15. ROSENBLUM, S. The effects of differential reinforcement and motivation on prediction responses in children. *Child Developm.*, 1956, 27, 99-108.
16. SCREVEN, C. G. The effects of interference on response strength. *J. comp. physiol. Psychol.*, 1954, 47, 140-144.
17. SEWELL, W. H., & MUSSEN, P. H. The effects of feeding, weaning, and scheduling procedures on childhood adjustment and the formation of oral symptoms. *Child Developm.*, 1952, 23, 185-191.
18. SHEPARD, W. O. The effects of verbal pretraining on discrimination learning in preschool children. Unpublished doctoral dissertation, State Univer. of Iowa, 1954.
19. SPENCE, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, 44, 430-444.
20. SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938. P. 138.
21. SPIKER, C. C. The effects of number of reinforcements on the strength of a generalized instrumental response. *Child Developm.*, 1956, 27, 37-44.
22. SPIKER, C. C. The stimulus generalization gradient as a function of the intensity of stimulus lights. *Child Developm.*, 1956, 27, 85-98.
23. TERRELL, G. The role of incentive in discrimination learning in children. *Child Developm.*, 1958, 29, 231-236.
24. TERRELL, G., & KENNEDY, W. A. Discrimination learning and transposition in children as a function of the nature of the reward. *J. exp. Psychol.*, 1957, 53, 257-260.
25. THURSTON, J. R., & MUSSEN, P. H. Infant feeding gratification and adult personality. *J. Pers.*, 1951, 19, 449-458.

SOCIALIZATION OF AGGRESSION AND THE PERCEPTION OF PARENTS IN FANTASY

JEROME KAGAN

The Fels Research Institute, Antioch College

Research on both the prediction and determinants of overt aggression in children has usually focused on the intensity of aggressive motivation as a major antecedent variable. For example, in studies attempting to predict overt aggression from related fantasy it has been usual practice to correlate a broad index of fantasy aggression with a category of overt aggressive behavior like fighting or delinquent activity. The implicit hypothesis seems to be that frequency of fantasy aggressive acts is an index of aggressive drive and should be positively correlated with occurrence of overt aggression. Most of these studies failed to find a positive and linear relation between these two variables (7, 11, 15, 16). However, recent investigators (5, 9, 12, 13) have stressed that aggressive motives are conflictful and subject to inhibitory influences and they have attempted to measure the strength of these avoidance or inhibition tendencies. In these more recent studies evaluation of fantasy indices of aggression anxiety significantly improved the power of the fantasy behavior to predict occurrence of overt aggression.

The present research utilized hypotheses concerning the child's acquisition of inhibitions on aggression to predict types of fantasy content that should be related to overt aggression in young boys. Current theorizing about the developmental determinants of aggressive behavior suggests that a perception of the parent as hostile and nongratifying influences the child's predisposition to aggressive behavior in two ways. It has been assumed that these conditions increase aggressive motivation and, in addition, are contrary to the conditions which motivate the learning of prohibitions on aggression. With respect to the latter, it is suggested that an important motive for the acquisition and practice of parental prohibitions on aggression is anxiety over anticipated loss of parental love. The more intense the anxiety over this anticipated loss the stronger will be the re-

sponse of adhering to the prohibitions of the parents. That is, the child trades the gratification derived from being aggressive for the continued love of his parents. Mahler (10) has stated,

Education, from the simplest gradual domestication and training of the infant up to the learning of skills and high scholastic achievement, can be compared with a continual barter in which the child is brought to give up infantile, egoistic and increasingly unacceptable, immature and objectionable behavior in return for tangible or intangible premiums, symbolizing love (p. 45).

It is hypothesized that the degree of anxiety generated by anticipation of loss of parental love depends, in part, on (a) the degree to which the child feels dependent on the parent for support in time of need and (b) the degree of nurturance given the child. These two variables are apt to be positively related and difficult to assess separately since a child should gradually extinguish on dependent overtures to parents if requests for support and nurturance were not gratified to some extent. Furthermore, an excessively nurturant and indulgent mother tends to prevent independent behavior from developing and the strength of the child's dependent tendencies remains at a high level. The position taken here is that the combination of a dependent child and a nurturant parent is most apt to generate anxiety over the anticipation of withdrawal of nurturance since the child feels he requires continued parental support. Thus, a dependent child who perceived the parents as nurturant would be highly motivated to adopt the prohibitions of the parents since disobedience should elicit anxiety over possible rejection. On the other hand, the child who was minimally dependent on the parent and who perceived the parent as hostile and nongratifying should feel less anxiety over anticipation of loss of love for such a loss would be viewed as less dangerous. Such a child would be minimally motivated to adopt and practice parental prohibitions. Existing data support these ideas. Several research reports suggest both a positive relation between maternal rejection and overt aggression in the child and an inverse relation between maternal overindulgence and the child's aggressive activity (1, 3, 8, 19). In a more recent study based on interviews with middle and working class mothers (17), Sears *et al.* found, for boys, a positive relation between dependency on the mother and a measure of the child's conscience and an inverse relation between maternal rejection and degree of conscience development. The authors conclude, "the pattern most calculated to produce 'high conscience' should be that of mothers who are usually warm and loving and then, as a method of control, threaten this affectionate relationship" (p. 388).

One final set of statements must be made explicit before the specific predictions can be stated. First, among young boys one can differentiate between unprovoked and justified aggression (8, 14). Although the former is generally punished by parents, the latter is often approved and encouraged by peers and family, and the young boy learns that aggression is

expected and justified when he is attacked by his social environment. If it is further acknowledged that most boys encounter situations in school justifying some form of aggressive retaliation, then chronic and consistent absence of aggressive behavior would indicate very strong inhibitory responses on overt aggression.

On the basis of the above discussion it was predicted that in comparing extremely aggressive and nonaggressive boys the former should perceive the parents as less gratifying and more hostile and show less dependent behavior toward the parents than the latter.

In this study the major technique used to measure the child's perception and relationship to his parents was a fantasy situation in which the child told stories to pictures specially devised to suggest themes of dependence on adults, nurturance from adults, anger between a child and parent, and parental punishment. It was expected that nonaggressive boys would report more dependency and parental nurturance themes and fewer themes involving anger between parent and child than a group of extremely aggressive boys.

A major problem in assessing the fantasy involved the question of whether certain fantasy content is or is not representative of the child's environmental behavior and experiences. Some investigators (15, 18) have implied that fantasy is apt to be representative of the child's environment and behavior when the fantasy behaviors and events are not goal states that are usually prohibited and/or frustrated. Parental anger and hostility are not usually regarded as desired goals, and dependent behavior toward an adult is not strongly prohibited in children. In this regard, Kagan and Mussen (6) found a significant positive relation between dependency themes on the TAT and overt dependent behavior among a group of male college students. It is therefore suggested that the specific fantasy content under study might be dominated by themes and actions representative of reality.

PROCEDURE

The subjects were 118 boys (Ss) drawn from seven classes (grades one through three) in a Columbus, Ohio, public school. Their ages ranged from 6-1 to 10-2 with a median age of 7-0. The school population was predominantly middle class, the majority of the fathers being skilled laborers and tradesmen. Before initiating interviews with each S the experimenter (E) spent several hours with each of the seven classes in order to become familiar with the children. After introductory games had established rapport in the interview, the E introduced the fantasy task with the following statement:

Let's play a guessing game now. I have some pictures here and I want you to guess what's happening on the picture. You have to make up a story to tell what you think is going on in the picture.

CHILD DEVELOPMENT

Thirteen pictures were administered, nine of which illustrated either a man or woman interacting with a small boy. The remaining four showed boy-boy interactions and are not relevant to this report. The nine adult-child pictures, in the order of their presentation, were as follows:

1. A boy is sitting on a chair holding a broken shoe lace and a woman is standing in the background.
2. A crying boy is sitting on the floor and a woman is standing behind the boy, looking down at him.
3. A boy with bowed head is standing near a bicycle which is turned over and a man is looking down at the boy.
4. A boy with bowed head is approaching a man sitting in a chair.
5. A boy is walking away from a woman who has her hands over her mouth.
6. A boy is holding his hand out to a man who is standing with one hand in his pocket.
7. A boy is crying and a woman is bending over the boy.
8. A boy with a cut on his arm is crying and a woman is standing in the background with her back to the boy.
9. A boy is crying and a man is standing in back of the boy looking down at him.

After the pictures were administered the *E* asked each *S* some questions, three of which are relevant for this report:

1. Who is the boss in your house, your mother or your father?
2. If your mother said one thing and your father said something different who would you listen to, your mother or your father?
3. Let's make believe you were bad at home and your mother and father were both home, who would punish you, your mother or your father?

Each *S* was rated by one person, his teacher, on a five-point scale for the following behaviors: (a) tendency to start fights at the slightest provocation and (b) tendency to hold in his anger and not to express it overtly. These two behaviors were selected in order to obtain a check on the consistency of the ratings. It was assumed that an *S* with a low rating on (a), i.e., never starts fights, would receive a high rating on (b), i.e., always holds in his anger. If the raters were consistent, there should be a high negative correlation between the ratings on these two variables. The product-moment correlation was $-.94$.

On the basis of the teacher ratings a group of 21 extremely aggressive boys (Group A) and 21 extremely nonaggressive boys (Group NA) were chosen from the entire sample. These two groups were composed of the three most and three least aggressive *Ss* in each of the seven classes. The *Ss* in both groups had been given extreme ratings on both variables.

Although the fantasy stories were scored for a large variety of themes, only the following content categories appeared with considerable frequency to the nine adult-child pictures.

1. *Dependency*—themes in which a child sought help from an adult with a problem situation.

2. *Nurturance*—themes in which a child was given unsolicited help with a problem; an adult showed concern for a child's welfare or a child was given money, food or gifts.

3. *Anger to a parent*—themes in which a child expressed anger toward a parent.

4. *Parental anger to a child*—themes in which a parent was angry with a child.

5. *Punishment*—themes in which a parent physically punished a child by spanking, slapping, hitting.

The themes were scored independently by the *E* and a graduate student in psychology without knowledge of the behavior ratings of the *Ss*. Percentage of agreement was high for each category and the over-all percentage of agreement for these five categories was 93 per cent.

RESULTS

Fantasy Data

1. *Dependency*. Occurrence of dependency themes was more frequent for the nonaggressive *Ss* with 9.5 per cent of Group A and 33.3 per cent of Group NA reporting one or more dependency themes ($p = .07$). All probability values are for one tail and were evaluated by the exact method described by Fisher (2).

2. *Nurturance*. There was no significant difference between the groups with respect to nurturance themes: 57.1 per cent of Group A and 61.9 per cent of Group NA reported this type of theme. However, one of the nurturance categories, "adult concern for a child's welfare," did tend to differentiate the groups with respect to the role of the person who showed this concern. The *Ss* in Group NA labeled the concerned adult as either a mother or father while the aggressive *Ss* described the concerned adult as a policeman, maid, stranger, or other nonparent figure. For the 15 *Ss* who told themes of adult concern, 28.6 per cent of Group A and 87.5 per cent of Group NA labeled the concerned adult as a parent ($p < .05$). One might infer from this finding that nonaggressive *Ss* were more likely than the aggressive ones to perceive the parents as nurturant figures. Supportive data from a study with older boys will be presented later.

3. *Parental anger*. As predicted, more of the *Ss* in Group A than in Group NA told one or more stories depicting parental anger toward a child (76.2 per cent versus 38.1 per cent; $p < .02$).

4. *Anger to parent*. The number of *Ss* in each group telling themes involving anger toward a parent also differentiated the two groups with 38.1 per cent of Group A and 9.5 per cent of Group NA telling one or more of these themes ($p < .05$).

5. *Punishment*. There was no significant difference between the groups on frequency of punishment themes: 52.4 per cent of each group reported one or more of these themes. It is possible that parental punishment is not a sensitive indicator of the gratifying or rejecting nature of the

CHILD DEVELOPMENT

parent-child relationship, for Hollenberg and Sperry (4) also found a very low correlation between ratings of maternal frustration and punitiveness on the basis of interview data.

Interview Data

1. *Child's perception of the boss of the family*. Two of the Ss in Group A had no father living with them and they were excluded from the statistical computations. To the question "Who is the boss in your house, your mother or your father?" more of the nonaggressive Ss answered "Mother" (42.9 per cent versus 10.5 per cent; $p < .05$).

2. *Child's preference for obedience*. To the question "If your mother said one thing and your father said something different who would you listen to, your mother or your father?" 57.1 per cent of Group NA and 26.3 per cent of Group A answered "Mother" ($p < .05$).

3. *More punitive parent*. To the question "Let's make believe you were bad at home and your mother and father were both at home, who would punish you, your mother or your father?" 47.6 per cent of the Ss in Group NA and 21.1 per cent of those in Group A viewed the mother as the major punitive agent ($p = .09$).

DISCUSSION

The aggressive boys produced less themes in which the child showed dependent behavior toward adults or a parent showed concern for a child's welfare and more themes in which the parent-child interaction was characterized by anger. These data tend to support the prediction of a correlation between overt aggressive behavior and fantasy content depicting a hostile parent-child interaction and weak dependent tendencies toward the parent. If the fantasy can be regarded as representative of reality, these results may be interpreted as partial validation for the hypothesis that prohibitions on aggressive behavior are more likely to be learned and practiced when the child is dependent on his parents and perceives them as nurturing and gratifying. Under these conditions the child should be anxious over possible loss of parental gratifications if he fails to adopt the parental prohibitions on aggressive behavior.

The answers to the three direct questions seem compatible with this interpretation. The nonaggressive boys were more likely to describe the mother as the major punitive agent, the parent they would be likely to obey, and "the boss" in the home. This perception of the mother as the dominant figure in the home suggests that the nonaggressive boys would be more likely than the aggressive ones to fear alienation from the mother and therefore more likely to adopt and practice her prohibitions on aggression. If it is further assumed that middle class mothers are apt to place stronger prohibitions on overt aggression than the fathers, the positive

relation between fear of alienation from the mother and absence of overt aggression tends to support the hypothesis of the paper. It is sometimes stated that one measure of degree of identification in a child is the degree to which prohibited motives have become socialized, and this process has sometimes been labeled "conscience" development. However, the child may use either his mother, father, or other figures as models for identification. The present relation between strong inhibition of aggression and a perception of the mother as authoritarian suggests that a measure of "conscience" development is not necessarily an index of intensity of identification with the same sex parent.

As mentioned earlier, data from a research project at a military school¹ furnished some support for the finding that when a fantasy stimulus suggested nurturance from an adult to a child figure, aggressive Ss were not apt to label the older figure as a parent despite the fact that this interpretation was quite congruent with the external stimulus.

From a total entering class of 683, 170 boys wrote stories to selected TAT cards. The dependent measure of overt aggression for this sample of older adolescent males (age range 17 to 22) was the number of demerits accrued by each boy for violations of regulations during the first five months at the school. The distribution of demerits for the entire class of 683 was divided into quartiles and 29 of the Ss who wrote TAT stories fell in the lowest demerit quartile (low on violations) and 29 were in the highest demerit quartile (high on violations). Analysis of the stories written to TAT card 7 BM which involved nurturance from the older to the younger man (63 per cent of the themes) revealed that 70.0 per cent of the low aggressive men and 31.3 per cent of the high aggressive men labeled the older nurturant figure as the father of the younger one ($p < .05$).

Additional data from a succeeding first year class at the same school furnish similar results. The entire class of 765 subjects wrote short stories to specially devised pictures, two of which illustrated an adolescent male and an older man. The first depicted a boy standing in front of a seated man and the second depicted an adolescent male sleeping in a bed with an older man leaning over the sleeping boy. Approximately one third of the group told stories to each picture in which the older figure nurtured the younger one. The demerit distribution for the entire class was standardized, divided into quartiles and each S assigned a score of 1 to 4 depending on the demerit quartile into which he fell. Comparison of the mean demerit score between the Ss telling themes of nurturance from a father versus themes of nurturance from a nonparent figure revealed that for each picture the latter group had a higher demerit score ($p < .15$). Although

¹ These data were obtained while the author was collaborating on the Medical Research Project, U. S. Army Hospital, West Point, New York, and the author wishes to thank Lt. Col. William Hausman for permission to report these data. The conclusions are those of the author and should not be construed as representing those of the Department of the Army.

these differences are not highly significant, the results are in the expected direction. One reason for the rather low significance value could be that both of these pictures were less suggestive of a nurturant relation between the adult and the boy than card 7 BM of the TAT. In any case, the results from these three independent studies all suggest that when a fantasy stimulus is congruent with an interpretation of a nurturant relation between a parent figure and younger male, failure to label the adult figure as a parent is positively correlated with overt aggressive behaviors. One might infer that failure to label the nurturant adult as a parent reflects a perception of the parent as nonnurturant.

Although this report has focused on the socialization of aggression, it is believed that anxiety over anticipated loss of nurturance acts as a motive for the socialization of other behaviors, e.g., sexuality and passivity, which also are subject to parental prohibitions. In addition to the partial validation of the hypotheses regarding the socialization process, a second implication of these data concerns the possible predictive power of these fantasy categories with respect to overt aggression. The findings tend to affirm the initial statement that fantasy content which is theoretically related to inhibition of aggression may be a more sensitive predictor of overt aggressive tendencies than the usual practice of summing the number of aggressive acts that a subject includes in his fantasy productions.

SUMMARY

It was predicted that the fantasy stories of extremely aggressive boys would contain more hostility between parent and child and less dependency on adults than the stories of nonaggressive boys. This prediction was based on the assumption that minimal dependency on parents and a perception of them as hostile would oppose the adoption and practice of parental prohibitions on aggressive behavior, while dependent tendencies and a perception of the parent as nurturant should facilitate socialization of aggressive motives. Individual interviews were held with 118 boys, ages 6-1 to 10-2, in which stories to pictures and answers to direct questions were obtained. Teachers rated the aggressive behavior of the children and the data for the 21 most aggressive and the 21 least aggressive boys were reported. The results showed that more of the nonaggressive boys produced themes of dependency on adults and more of the aggressive boys told stories involving anger between parent and child. In answers to direct questions the nonaggressive children perceived the mother rather than the father as the major punitive agent and the parent they would be most apt to obey. This result indicated that the nonaggressive boys were more anxious about alienation from the mother than from the father. It was suggested that anxiety over alienation from the mother should lead to relatively strong inhibitory responses with respect to aggressive behavior since mothers are apt to be less permissive of aggression than fathers.

REFERENCES

1. BALDWIN, A. L., KALHORN, JOAN, & BREESE, FAY H. Patterns of parent behavior. *Psychol. Monogr.*, 1945, 58, No. 3.
2. FISHER, R. A. *Statistical methods for research workers*. (7th Ed.) Edinburgh: Oliver & Boyd, 1938.
3. HATTWICK, B. W. Interrelations between the pre-school child's behavior and certain factors in the home. *Child Developm.*, 1936, 7, 200-226.
4. HOLLENBERG, ELEANOR, & SPERRY, MARGARET. Some antecedents of aggression and effects of frustration on doll play. *Pers. Sympos. Topical Issues*, 1951, 1, 32-43.
5. KAGAN, J. The measurement of overt aggression from fantasy. *J. abnorm. soc. Psychol.*, 1956, 52, 390-393.
6. KAGAN, J., & MUSSEN, P. H. Dependency themes on the TAT and group conformity. *J. consult. Psychol.*, 1956, 20, 29-32.
7. KORNER, ANNELIESE F. *Some aspects of hostility in young children*. New York: Grune & Stratton, 1949.
8. LESSER, G. S. Maternal attitudes and practices and the aggressive behavior of children. Unpublished doctor's dissertation, Yale Univer., 1952.
9. LESSER, G. S. Conflict analysis of fantasy aggression. *J. Pers.*, 1958, 26, 29-41.
10. MAHLER, M. S. Ego psychology applied to behavior problems. In N. D. C. Lewis & B. C. Pacella (Eds.), *Modern trends in child psychiatry*. New York: International Universities Press, 1945. Pp. 43-56.
11. MURRAY, H. A. *Thematic apperception test manual*. Cambridge: Harvard Univer. Press, 1943.
12. MUSSEN, P. H., & NAYLOR, H. K. The relationships between overt and fantasy aggression. *J. abnorm. soc. Psychol.*, 1954, 40, 235-240.
13. PITTLUCK, PATRICIA. The relation between aggressive fantasy and overt behavior. Unpublished doctor's dissertation, Yale Univer., 1950.
14. RADKE, MARIAN J. The relation of parental authority to children's behavior and attitudes. *Univer. Minn. Inst. Child Welf. Monogr. Series*, 1946, 22.
15. SANFORD, R. N., ADKINS, M. M., MILLER, R. B., COBB, E. A., & others. Physique, personality and scholarship: a cooperative study of school children. *Monogr. Soc. Res. Child Developm.*, 1943, 8, No. 1.

CHILD DEVELOPMENT

16. SEARS, R. R. Relation of fantasy aggression to interpersonal aggression. *Child Develpm.*, 1950, 21, 5-6.
17. SEARS, R. R., MACCOBY, ELEANOR E., & LEVIN, H. *Patterns of child rearing*. Evanston, Ill.: Row, Peterson, 1957.
18. TOMKINS, S. S. The present status of the thematic apperception test. *Amer. J. Orthopsychiat.*, 1949, 19, 358-362.
19. UPDEGRAFF, RUTH. Recent approaches to the study of the preschool child. *J. consult. Psychol.*, 1939, 3, 34-36.

CHILD DEVELOPMENT

Vol. 29

September, 1958

No. 3

Editor

WILLIAM E. MARTIN
Purdue University

Consulting Editors

BETTYE McDONALD CALDWELL
Washington University School of Medicine

ERCEL S. EPPRIGHT
Iowa State College

EUGENE S. GOLLIN
Queens College

HARRY LEVIN
Cornell University

PAUL H. MUSSEN
University of California

KATHERINE E. ROBERTS
Merrill-Palmer School

WATARU W. SUTOW
University of Texas

WILLIAM W. WATTENBERG
Wayne State University

Managing Editor

KATE HOFFMAN

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS
(of The Society for Research in Child Development, Inc.)
Purdue University
Lafayette, Indiana

Made in United States of America

Notice to Contributors

Child Development is devoted to basic research and theory concerning the growth and development of the child.

Manuscripts should be submitted in duplicate, each copy bearing only the title as a means of identification. The title, author's name, and institutional affiliation (together with any footnotes which reveal the identification of the author) should be submitted on separate pages.

Contributors are asked to refer to the *Publication Manual of the American Psychological Association* (1957 Revision) in the preparation of manuscripts.

The following rules require special emphasis:

1. All copy, including indented matter, should be typed double-spaced on white standard paper. Lines should not exceed six inches in length.
2. Footnotes should be few in number. They should be indicated in the text by superscript arabic numerals, numbered consecutively throughout the manuscript. The footnotes themselves should be typed double-spaced on a separate page placed at the end of the manuscript.
3. Tables and figures are costly to print and should be used only when essential. The author will be asked to contribute to the cost of composition of tabular and graphic material, which is excessive in amount or unduly complicated in form. Each table should be typed on a separate sheet. Insert a guide line (e.g., Table 1 about here) at the appropriate place in the manuscript. Figures should be prepared with black India ink, suitable for photography. Legends typewritten on an ordinary typewriter are unacceptable.
4. References in the text should be indicated by numbers in parentheses that refer to the list of references at the end of the article. Do not use footnotes for this purpose. In the list of references, the arrangement should be alphabetical by author's surname, or for multiple authors by the name of the senior; for any one author, by order of date. The following standard form should be observed:

Jones, H. E. Development in adolescence. New York: D. Appleton-Century, 1943.

Shepard, Winifred O., & Schaeffer, M. The effect of concept knowledge on discrimination learning. Child Development, 1956, 27, 173-178.

The World List of abbreviations is used for journals.

The manuscript submitted should represent the final form in which the author wishes the paper to appear. All changes in proof (other than printer's errors) will be charged to the author.

All manuscripts should be addressed to William E. Martin, Editor, Child Development Publications, Purdue University, Lafayette, Indiana.

RETROSPECTIVE ATTITUDE STUDIES OF PARENT-CHILD RELATIONS

RICHARD Q. BELL

National Institute of Mental Health¹

Many reviews of the literature and considerations of methodological problems in specific areas are published after considerable experimentation and at a time when interest in the field has lagged. This paper will attempt to anticipate problems of method in a newly developing research approach in order to contribute to maximal use of research energy in attacking the problems involved. The approach consists of administering attitude questionnaires to parents whose children differ in behavior. In studies to be considered it is indicated or implied that variation in the independent variable as measured by parental attitude measures has contributed in mediated or direct form to variation in the dependent variable or criterion measure of child behavior. In further discussion, the term criterion variable will be used to refer to measures of the behavior of children, whether such measures result from the use of experimental and control groups, or from measurements of a continuous variable in a single group.

Studies to be discussed are *postdictive* rather than *predictive*. The study design thus parallels the retrospective clinical case study. It might be called a "reverse leap-frog" design. Inferences from obtained differences in parental attitudes involve a backward extrapolation in time, conceptually "leaping over" the actual behavior of parents with the children. The nature of an earlier developmental process is inferred from data obtained at a later point in time. Accordingly it is difficult to assert that obtained differences between parental attitudes measured after the development of children reflect factors operating at the time the children were developing. On the other hand, obtained differences in attitudes between parent groups cannot be ignored. It is tempting simply to suggest longitudinal studies of the effects of measured parental attitudes on child development. However, the productivity of such cumbersome, slow-moving projects may be heavily dependent on the preliminary hypothesis development carried out most rapidly in the retrospective type of study.

Eleven published studies of the retrospective type are listed in one review of the literature (34). A limited informal survey of more recent

¹ National Institutes of Health, Public Health Service, U. S. Department of Health, Education and Welfare.

CHILD DEVELOPMENT

studies involving parental attitudes revealed that the majority are retrospective and the volume considerable, 14 completed but not yet published and 62 planned or underway in 1957. Since the retrospective approach will have continuing importance as a source of new hypotheses in the field, it seems worthwhile to consider the circumstances which are likely to affect its contribution.

Methodology will be considered in the context of the substantive task of making inferences from the general type of data provided by the research design. The following problems will be considered: the problem of the influence of social context on attitude expressions, further problems which develop from attempts to cope with social context factors by item phrasing, problems of social context which indicate a need for special experimental populations and situations, problems emanating from the fact that attitudes may change over time, the problem of attitudes specific to a given child, and problems of linking attitudes to behavior. Since available texts on research design and statistics provide ample discussion of such topics as cross-validation and how to control the effects of basic demographic variables, this paper will be limited to less obvious difficulties which are generated by this specific research approach. Discussion will also be limited to the type of retrospective and cross-sectional studies which are within the capabilities of most investigators working in this area. Longitudinal approaches will be referred to only where necessary to an adequate discussion of issues involved.

PROBLEMS OF SOCIAL CONTEXT

Studies of factors affecting conformity to social pressures and group norms have in general indicated that different measured attitudes can be elicited from the same subjects in different social situations and that an attitude can be modified by different relationships between a respondent and the individual to whom he is expressing an attitude (23, 26, 36). Gordon (16) has reported in the case of personality test items that preference for socially acceptable items is a very powerful factor even in a situation where no question of securing a job, promotion, or admission to a professional school was involved. That this general tendency to respond in what is perceived as the socially acceptable fashion is almost automatic and unthinking is further suggested by the studies of Ash and Abramson (1) on college students and Stouffer (40, p. 504) on inductees which show little or no effect of anonymity on attitude questionnaire response. Parental attitude measures are probably no exception. Since they may be altered readily by appropriate films or discussions (39), it seems likely that they may also be susceptible to social context factors. Freeman and Grayson (14) noted that many mothers of adult schizophrenics requested permission to take parental attitude questionnaires home with them in order to obtain help from relatives. This reaction in a literate parent sug-

gests, among other things, a desire to say the "right thing" in a testing situation. Sophistication in current child rearing theories is likely to contribute in some fashion to measured response as is shown in a mean correlation of $-.27$ between educational level and tendency to agree with the content of a 36-item questionnaire consisting largely of statements in opposition to such theories (34).

An investigator who has obtained significant differences in parental attitude scores between two well matched groups whose children show adjustment differences may maintain that the differences must not be due to social context factors since the social context of testing was presumably the same for both groups. This seems unlikely where parents of children with deviant adjustment constitute one of the groups since it assumes that the parent does not know that his child is deviant or will not connect the fact with the situation in which he is being asked to fill out a questionnaire dealing with child rearing practices. It can be maintained, however, that the difference between groups may be a reflection of a differential response to the situation in which the questionnaire was administered, and that this difference in response to a social situation is itself significant. Goldstein and Carr (15) suggest that the refusal or inability to answer a questionnaire, as shown to a significant extent by mothers of adult male catatonics, reflects a personality trait possibly important in the etiology of their son's personality development. In this case, however, the questionnaire and its content become less important than the experimental situation created by the attempt to administer a questionnaire. The meaning of the social context to the respondent thus becomes even more important when the investigator intends to interpret obtained differences in terms of social context.

Respondents may be unable to answer questionnaires in a way that is indicative of their spontaneous, day-to-day behavior with children for reasons other than social context. Investigators postulating unconscious factors in the perception of the meaning of attitude items or defense mechanisms which operate regardless of variations in the social context may object to grouping such factors under the heading of social context. The grouping is for expository purposes and such distinctions would not be germane to the discussion since efforts to cope with both types of variables have been largely the same.

ITEM-PHRASING SOLUTIONS TO THE PROBLEM OF SOCIAL CONTEXT

According to a content analysis (34) the more recent parental attitude questionnaires show a definite tendency to cope with the problem of social context primarily by providing the parent with items having content in opposition to current theories. Items with this type of content tended to discriminate between criterion groups better than items affirming such theories or using mere statements of fact. Items affirming existing theories

tended to have very low dispersions, the knowledge of such theories apparently being sufficiently general that most parents tend to agree, rather than vary, in choice of response options.

To reduce expressions of disagreement by parents the content in opposition to current theories is disguised by phrasing which makes the item either slightly vague, apparently aphoristic, or somewhat attractive by virtue of rationalizations provided. Examples of such items from Mark's study (30) are shown below. As is typical in the more recent questionnaires, the subject is to respond in categories varying from strong agreement to strong disagreement.

A good mother should shelter her child from life's little difficulties.
Too much affection will make a child a softie.

A devoted mother has no time for social life.

Examples of items which failed to achieve high discrimination and were rated as stated in agreement with current theories are as follows:

Childhood is the foundation period of a person's life.

A child should have freedom to develop a personality of his own.

If socially acceptable content may be equated with content stated in agreement with current theories, current item writing approaches in the field of parental attitudes may be placed in the context of findings relative to item phrasing in the broader field of attitude measurement. Rundquist (33) has summarized a large number of attitude studies showing that items with socially unacceptable content are more effective discriminators between criterion groups across a broad spectrum of age groups, types of item presentation, and test format. Gordon (16) has reported that the more unclear a personality test item which contains socially unacceptable content the more valid it is likely to be. Disguised or projective item phrasing may thus increase the discriminating power of the items containing socially unacceptable content. The latter are effective in and of themselves but may possibly be enhanced by vague or deceptive wording. For the purposes of this paper it will be sufficient to note that the tendency toward this type of phrasing and content constitutes one of the major attempts to cope with the problem of social context.

PROBLEMS RAISED BY USE OF VAGUE OR DEVIOUS UNACCEPTABLE ITEMS

Interpretation of a single item is precarious. Basically, it is comparable to attempts to interpret a factor showing loading on only one test. McNemar (28) pointed out in criticizing single item opinion polling that percentage of agreement with a single opinion question can change drastically with minor changes in wording. Interpretation becomes still more hazardous when, as reflected in the above trend in phrasing parental attitude items,

we add the problem of determining whether the aphoristic quality of an item, its vagueness, the extent to which it embodies rationalizations, its lack of social acceptability, or some other factor is crucial. Interpretation of such items borders on interpretation of responses to ink blots. And yet, hypothesis development has had to be based on such interpretations of single items (14, 21) in the absence of more secure data. Accordingly, it is especially important for an investigator writing projective or vague items to define his concept with a demonstration of internal consistency in a group of items. Campbell (10) and Green (17) have indicated the logical necessity of such demonstrations in keeping with existing definitions of attitudes in terms of response consistency. Such scales or pools of homogeneous items should preferably contain items using different phrasing techniques. Scales for which internal consistency have been demonstrated have been reported but have not as yet been used extensively in retrospective studies (34, 44). Such scales at least render more plausible an investigator's efforts to link a concept with a measured attitude. Without such plausibility there is little interest in a failure to find results in the direction predicted from the hypotheses. The defect could exist at the "spade work" level of item writing rather than in the basic concepts.

A problem posed especially by high saturation of questionnaires with socially unacceptable content and the use of vague, ambiguous item phrasing is that of measuring response sets such as acquiescence, evasiveness, speed and guessing. Cronbach (11) has summarized evidence to indicate that such sets, which are most obtrusive in precisely the fixed category response situation used almost exclusively in parental attitude questionnaires, are maximized by stimulus situations which are unclear or ambiguous to the subject, are likely to change with different item content, and may reduce validity unless partialled out and used expressly to predict a criterion. Berg and Rapaport (6) have shown that response sets are present even where an imaginary questionnaire is administered under circumstances resembling an ESP experiment. Bass (2) has presented evidence to support the contention that three-fourths of the reliable variance in the F scale used extensively in studies of the authoritarian personality may be explained as an acquiescence set, only one-fourth being related to a content factor of authoritarianism. His finding is consistent with Stogdill's (38) report that there is a tendency in many parents to agree with attitude items having quite contradictory content.

Several possibilities exist for dealing with the problem of response sets in parental attitude questionnaires. Constructing items with directly contradictory content provides a check on the acquiescence set. This approach necessitates almost doubling the length of existing questionnaires and thus reduces the number of concepts which can be included. Another approach involves using forced choice formats, which have been found relatively free of response sets. In an informal experiment the author found that forced choice formats tended to generate much more serious rapport prob-

lems than those encountered with customary formats (discussed below). Another approach would permit use of existing questionnaires which provide several scores each based on a scale or pool of items (34, 44): the score for each scale for a given individual is expressed as a function of the individual's average and standard deviation for all scales. This procedure is described by Cronbach and Gleser (12) as a means of partialling level and dispersion out of profile measures to provide a pure measure of profile shape. The average of scores on all scales as described provides a measure of acquiescence and the standard deviation of scores on all scales a measure of variability for each subject.

A procedure such as this, which makes it possible to partial out but not preclude the operation of response sets, would seem desirable at the present, since measures of response sets in parents may be related to criterion measures of child behavior. This is an open question at present. Freeman and Grayson (14) reported that extremity and consistency of response scores failed to differentiate between two parent groups. On the other hand, Goldstein and Carr (15) report that a phenomenon very similar to response set, namely failure to answer questions, made it possible to distinguish between their criterion groups when responses to the parental attitude items as such did not. Investigators such as Guttman (18) and Elinson (13) have reported in the broader field of attitude measurement that direction of response may not be as important for many purposes as generalized intensity, the strength with which attitudes are manifested regardless of content. Further, many investigators use parental attitude scales as measures of personality (19, 21, 37) and thus would view with interest the relationship between evasiveness and tendency to check the response category "undecided" as mentioned by Cronbach (11). Also pertinent are the findings of Berg and Collier (5) that sets to accept extreme response categories in the Perceptual Reaction Test are related to personality measures such as confidence, suggestibility, and anxiety.

Investigators obtaining measures of response sets from existing questionnaires should keep in mind that sets such as acquiescence and variability may be specific to the kind of item being used. A parent conversant with the current popular literature on child rearing practices might show different response sets to items stated in agreement with, and items stated in opposition to, current theories. Berg (4) has reported that there is only a moderate correlation between response sets to verbal and spatial items in the Perceptual Reaction Test. The extent of variation of response sets relative to different kinds of parental attitude questionnaire content is not known at present.

In addition to the problem of response sets, concentration on socially unacceptable content in parental attitude items generates problems at the level of both theory and rapport with the subjects. One study has reported the necessity of using "rapport" scales in order to raise the level of satisfaction of sophisticated respondents who apparently did not feel that they could

adequately express their attitudes through responding to the content provided (34). Most serious is a probable loss in test efficiency reflected by lack of rapport with more sophisticated subjects. This may imply that the item preference values obtainable with the type of content widely used in parental attitude scales are poorly adjusted to the task of making discriminations between sophisticated subjects varying in true attitudes toward child rearing. It appears most likely that a desirable median difficulty level of .50 does not exist for many such questionnaires used with subjects at or above the high school level of education. Thus, problems of differentiating parent groups whose children show superior social or intellectual achievements will be difficult to pursue if a more adequate understanding of the success of the socially unacceptable items is not reached. For adequate exploration of the domain of parental attitudes it is important, as Kuder (25) has pointed out in the case of interest and personality inventories, that it be possible to develop, without distortion, a fairly reliable measure in any direction within the factor space represented. The need for experimental studies of item behavior similar to those carried out by Rundquist (33) and Gordon (16) seems acute in the field of parental attitude measurement. Such experimentation should focus among other things on the possibility of obtaining successful discrimination using concepts stated in agreement with existing concepts of child rearing and varying such factors as extremity of statement. Variation in social acceptance values and the accuracy of perception of current theories of child rearing (32, p. 414) should be studied further in different social classes.

EXPERIMENTAL DESIGN SOLUTIONS TO PROBLEMS OF SOCIAL CONTEXT

A special problem of social context arises from the fact that many parents tested in retrospective parental attitude studies may have been reacting to the very fact that their children's behavior has been socially identified as superior or unusual, whether leadership or adjustment problems are involved. The attitude pattern which differentiates them from a control group may then include reflections of factors such as guilt or self-aggrandizement rather than a pattern linked to the developmental history of the child. Fortunately, this problem may be attacked fruitfully in at least two different ways. The investigator minimizes the influence of this factor by using attitude measures from parents whose children have not been socially identified as having adjustment problems, under circumstances where it appears unlikely that the parent is aware of the child's status. Studies in the nursery school setting seem to meet this requirement (31). In general, the possibilities for studies of this type in the nursery school setting appear promising since extensive ratings of child behavior are possible.

Opportunities for partialling out self-justifying attitudes also exist through selection of special control groups while retaining the basic retro-

CHILD DEVELOPMENT

spective study design. McFarland (27) in a replication of Mark's study (30) obtained a control group of mothers of nonschizophrenic individuals by requesting adult males receiving hospital treatment for physical ailments to contact their mothers for participation in the study. Mark obtained some members of his control group in this fashion, but many more from women's clubs and other organizations. McFarland could contrast the attitude patterns of two groups of mothers both having adult sons with health problems sufficiently severe to require hospitalization. This offers an improvement logically over the Mark design since both groups of mothers were presumably reacting at least to the fact of hospitalization of their sons.

The control group used by McFarland still does not permit a control over possible differential reaction to the fact of a son's mental illness. Contrasting parents of children showing different mental illnesses, however, permits a reduction in the possible contribution of this gross factor to obtained differences between criterion groups (15). There are many possibilities for control groups which permit control over guilt and other "reactive" effects. Contrasts may be made between mothers of children showing mental retardation, brain damage, somatic disorders, psychosomatic disorders, autism, or impulse control problems as the problem dictates. No one control group meets all logical requirements, but each permits a better approximation of differential attitude patterns which may have contributed in a primary way to the direction of development of the child.

ATTITUDE CHANGE OVER TIME

If investigators accept the notion that parental attitudes are changed as a result of living with children, or simply show spontaneous change over time, it becomes very difficult logically to relate developmental events in time segment A-B of the child's life to parental attitudes measured at time B or C. This problem is minimized, again, by carrying out studies on parents of children at lower age levels, particularly in the nursery school setting. However, many important adjustment characteristics do not emerge until later in life and there are no existing ways of linking early adjustment with such later characteristics. The investigator must either study the phenomenon retrospectively in its later form or not study it at all, there being no very secure means of selecting a group of parents whose children later follow-up, even if patience existed for the task of follow-up.

Two possible approaches are suggested. One is to isolate parental attitude domains which are relatively stable in different age cross-sections and restrict attitude measurement to these. This will be very difficult to do in the general field of attitudes. The Allport-Vernon Scale of Values and the Strong Vocational Interest Inventory showed much greater consistency over time than did several attitude scales in Kelly's (22) follow-up study of married couples. Attitudes toward rearing children showed 20-year retest

correlations of .15. Even assuming for purposes of discussion that there has been some improvement in tapping more stable child rearing attitudes in parents, an investigator primarily interested in working with measures which show stability over time may well consider the possibility of using interest and value measures obtained from parents, considering the impressive test-retest correlations of such measures from the same study by Kelly.

Although it would seem reasonable to assume that psychological growth, a cumulative process, would be most influenced by stable attitude structures in the parents, some reasons exist for considering an alternative viewpoint. Tietz (43) and others have pointed out the possibility that there may be "age-specific" parents, parents who meet the needs of children during one period of life but fail to assist their growth in some later phase. Such individuals may in effect lack the flexibility to grow as a parent along with the child. It seems likely that other parents may show important attitude shifts over time but change predictably, so that the child has been exposed to an attitude channel over time, each point belonging to an over-all component process. It may be that there is a natural history of attitude change in parents with different kinds of personality structures and that knowing these attitude changes an attitude channel can be estimated, much as has been done for the prediction of physical growth. According to this view, attitude differences between two groups of parents could be considered meaningfully related to the adjustment of the children, even though such differences would only represent terminal phases of the process which had changed over time. The problem would then be to locate and identify the early and in-between points in the process. Approximations to these attitude changes over time might be obtained by cross-sectional or accelerated longitudinal methods (3) prior to full scale longitudinal studies.

PARENT-CHILD ATTITUDE SPECIFICITY

It seems almost unnecessary to point out that parents are capable of reacting quite differently to various children in the same family. Reactions to such a gross factor as sex differences have been noted (20). Favoritism is very frequently observed in families as well as many other special roles accorded to children. Thomas (42), Tietze (43), and many others have asserted that mothers of schizophrenic individuals had quite different relationships with their nonschizophrenic children from an early point in time. The type of retrospective study discussed in this paper very frequently involves parents who have two or more children, and yet the instrument which measures the parents' attitudes does not provide any indication of the particular child to whom the attitudes refer. This failure to determine the referent of the attitude would seem to lead to at least two major consequences. First of all, differences between criterion groups will be obscured. The attitudes of mothers in groups being contrasted may reflect orientations

toward a child other than the particular child from the family whose behavior led to assignment of the mother to a given criterion group. Investigators who obtain significant differences between groups may maintain that their results were obtained in spite of rather than because of this factor. On the other hand, many more differences might have emerged had this factor been controlled and the study is correspondingly lacking in precision. Secondly, preoccupation with this type of study neglects the fact that an adequate assessment of the general impact of parents on all of the children in their family has not been carried out. The particular child selected from the family for study may provide misleading information on the behavior characteristics common to all children in the family and differentiating the family from other families.

To overcome the logical absurdity of having nonspecific parental attitude measures in a situation where the criterion measure invariably involves a single child from a family, certain measures of parental orientations not now found in attitude scales may be necessary. Scales for measuring differences in attitudes toward rearing female and male children may be useful. New variance may be brought into the format by sets to respond in accordance with attitudes toward specific children. Parents who vary greatly in orientations toward different children may be manifesting either a laudable sensitivity to the needs of different children, or be overinvolved and working through internal conflicts with a particular child. Parents who have reared two or more disturbed children but with qualitatively different relationships could be contrasted attitudinally with suitable control groups to study this problem retrospectively. It should be noted, however, that a testing procedure which links attitude expressions in the parent's mind to specific children has some likelihood of marshalling self-justifying attitudes. It may be desirable to use sets to respond in terms of different sexes, ages, and personalities of children to help identify the probable referent of their attitude expressions without eliciting self-justifying attitudes. Another possibility is to make studies of one-child families or of very young families prior to the birth of a second child.

Closely related to the subject of parent-child attitude specificity is the question of the feed-back effects on the parent of living with a particular kind of child. At first glance this question would seem to require longitudinal studies. It is noteworthy that interesting data on this question have been provided by retrospective studies in which parents of children with congenitally related behavior problems are contrasted with parents of normal children. Mann (29) administered parental attitude questionnaires to mothers of children with cerebral palsy and a closely matched group of mothers of normal children. Many of the large number of significant differences found between the groups pointed to the possible operation of a reactive factor in parental attitudes when measured retrospectively. Apparently the reaction of a parent to the problems of child rearing posed by a handicapped child may lead to manifest attitudes resembling those which when

manifested by the mother of a schizophrenic or delinquent child have been assumed to indicate causal significance (21). Klebanoff (24) has reported that it is possible to obtain significant differences in attitudes between mothers of normal children and mothers of schizophrenic children, or between mothers of normal children and mothers of mentally retarded or brain damaged children. It is not possible in his study, however, to obtain significant differences between mothers of schizophrenic children and mothers of mentally retarded or brain damaged children. These studies suggest a number of changes in the attitudes of the mothers themselves, in the nature of the marital relationship, and in the relationship of the home to the community which may be expected to result from rearing a congenitally handicapped child. The methodological significance of the above studies lies in the fact that they provide findings which serve as a check on the reports of other retrospective studies which have not taken feed-back factors into consideration. Longitudinal studies of concepts such as the "schizophrenogenic mother" would clearly seem premature when data even from retrospective studies cast doubt on the basic formulations.

Further clarification of feed-back effects may be expected by contrasting parents of normal children with parents of children showing other behavior problems linked to congenital disorders. This approach need not be limited to behavior problems linked to congenital factors, however. Contrasts may also be made between attitudes of parents whose adopted children showed differences in behavior prior to adoption.

DIRECT STUDIES OF ATTITUDE-BEHAVIOR LINKAGE

In order to clarify further relations between parental attitudes and behavior of children it would seem vital to study one major link in the process, that of the relation between parental attitudes and what parents actually do with children. An apparently simple and straightforward approach to this problem would be to measure the attitudes of parents and then study their behavior with children, preferably in the home, but possibly also in such settings as cooperative nursery schools and neighborhoods. This type of study can be carried out just as readily as a retrospective study, as far as time considerations are concerned, since both involve obtaining measurements in a single cross-section of time. Accordingly, it is surprising at first glance that no direct studies of relations between measured parental attitudes and actual behavior of parents with children can be located in the literature. This lack of studies, however, seems much more understandable when one explores in more detail just one example of the problem of obtaining even reasonably adequate observations of parent behavior. It is immediately obvious in the case of observations in the home that it will be difficult to penetrate social facade, and that limited behavior samples may not bring low frequency but highly significant critical incidents under observation. Most home visits are not made during two periods

of the day when common sense observations would suggest that a large number of critical incidents occur. These are the periods immediately prior to dinner and bedtime for the children. Home visits as customarily limited in number and appropriateness of timing provide useful estimates of the child rearing climate of the home, but perhaps no better estimates than an interview with the parents in some setting other than the home. The home remains for all practical purposes a "little black box"; parental attitude measures, home visits, and interviews with parents and children represent some ways of obtaining estimates of its internal processes; the behavior of children in community facilities is a readily observable output. It is easy to observe children but difficult to observe parents and children in the home.

The importance of the above type of study should not be minimized in view of the difficulties. On the other hand, there should be some consideration of a more fragmentary approach to the same problem, studies in which parental attitudes are related to actual behavior of a parent with other adults. There are several reasons why this kind of study may assist the development of the field: spontaneous moment-to-moment behavior of parents with children in the home is relatively inaccessible; the complexity of attitude structure and its linkage to behavior requires comprehensive studies which current assessment programs worked out for adults can provide; finally, in view of the apparent instability of measured attitudes cited in the study by Kelly (22) it may become increasingly important to find ways of linking verbalized attitudes to personality traits which have stability. In this way a mediating set of propositions linking attitudes to adult personality may be developed. This in itself is an ambitious objective from a theoretical standpoint, without regard for the next necessary step of providing a set of propositions which will mediate between such a network and actual behavior with children.

The complexity of relationships in this area is suggested by the fact that one factor analysis of 24 attitude measures in a very homogeneous population of unmarried females resulted in the identification of five orthogonal factors (35). In the general field of attitude measurement Bray (9) has reported complex interactions between attitude and personality scores and actual behavior of adults in a situation involving interracial behavior. Block (7), reporting data obtained from a comprehensive assessment of a population of male army officers who were also parents, found that expressed parental attitudes of restrictiveness were related to ratings of submissiveness, suggestibility, conformity, indecisiveness, and ineffectuality received by the individuals when observed in interactions with other adults. These studies suggest that a simple isomorphic projection of attitude concepts into behavior will be found only occasionally. Not enough is known about the structure of attitudes in this area with different population groups to permit an investigator to single out one attitude, test for a single behavioral referent, and attribute significance to a negative result

in the event that attitude-behavior correspondence is not found. This type of result will only be of interest to investigators who take attitude scale labels very seriously and regard behavior as simply linked thereto. Again, as in the problem of interpreting the meaning of responses to items, loss of research energy may simply result from negative results obtained in studies which have defined attitude or behavior variables without a convincingly adequate battery of measures. Adequate sampling of both domains need not imply unimaginative "fishing expeditions." Experimental procedures merely need to be linked to independent and dependent variables which are defined in multivariate terms. This is well within the possibilities of existing techniques such as analysis of variance designs for testing pattern differences between groups (8). Assessment programs for adults provide the type of data which are required, with the addition of parental attitude measures to testing batteries. There are, of course, many other populations of adults made accessible to behavior studies but not seen at present as having additional psychological importance because they are also parents.

SUMMARY

This paper has reviewed problems of method emanating from current use of attitude questionnaires in a rapidly developing field of investigation typically involving testing for attitude differences between parents of children showing different kinds of behavior. The retrospective design offers a means of generating new hypotheses and carrying out preliminary tests prior to long term studies. The design provides equivocal data on developmental hypotheses due to the fact that parental attitudes as manifest may be subject to social context factors such as parental needs for self-justification, may have changed over time, may have behavioral impact on different children in quite different ways, and may be linked to actual behavior in very complex fashion. Many of these problems can be handled by refinements in the retrospective design itself.

Recent use of the projective item to cope with social context factors creates problems in that the items are more difficult to interpret and response sets maximized. Demonstration of internal consistency becomes especially important. Response sets which are maximized by such items in the usual formats may be controlled or partialled out by using new scoring techniques on existing questionnaires, writing scales with contradictory content, or using forced choice formats. The use of items which are vague and stated in opposition to socially acceptable or expert views on child rearing poses problems in rapport and discrimination among subjects at higher educational levels. Basic research on item behavior is important to further development of this approach.

Self-justifying attitudes in the respondent can be minimized by using parents of children who vary in behavior but have not been socially identified, either by stigma or praise, as showing unusual adjustment. Various

CHILD DEVELOPMENT

control groups can be used which permit partialling out specific kinds of self-justifying attitudes likely in parents of deviant children.

Lack of stability of measured parental attitudes over time is a less critical problem of studying adjustment of young children. Studies of stable attitude areas or the use of other measurement approaches may be necessary to provide meaningful measures of parental orientations to which a child is exposed over long time periods.

The problem of parent-child attitude specificity need not be considered by an investigator interested primarily in generalized family climate, but it is a crucial problem for prediction and for the development of a more sharply focused theory. Special scales and instructions to the respondent may be needed to provide data in this area.

No direct studies have been carried out on the relation between parental attitudes measured by questionnaires and the actual behavior of parents with children. Pending solution of current problems in obtaining accurate information on spontaneous day-to-day behavior of parents with children, studies of the relation of parental attitudes to behavior with other adults may help develop an intermediate set of propositions linking parental attitudes and parental personality. Such studies may assist in ordering the complexity of problems involved in linking attitudes to actual behavior with children, and may provide data on more stable characteristics of parents which conceivably have long term cumulative effects on the behavior of children.

REFERENCES

1. ASH, P., & ABRAMSON, E. The effect of anonymity of attitude questionnaire response. *J. abnorm. soc. Psychol.*, 1952, 47, 722-723.
2. BASS, B. M. Authoritarianism or acquiescence? *J. abnorm. soc. Psychol.*, 1955, 51, 616-623.
3. BELL, R. Q. An experimental test of the accelerated longitudinal approach. *Child Developm.*, 1953, 25, 281-286.
4. BERG, I. A. The reliability of extreme position response sets in two tests. *J. Psychol.*, 1953, 36, 3-9.
5. BERG, I. A., & COLLIER, J. S. Personality and group differences in extreme response sets. *Educ. psychol. Measmt.*, 1953, 13, 164-169.
6. BERG, I. A., & RAPAPORT, G. M. Response bias in an unstructured questionnaire. *J. Psychol.*, 1954, 38, 475-481.
7. BLOCK, J. Personality characteristics associated with father's attitudes toward child-rearing. *Child Developm.*, 1955, 26, 41-48.
8. BLOCK, J., LEVINE, L., & McNEMAR, Q. Testing for the existence of psychometric patterns. *J. abnorm. soc. Psychol.*, 1951, 46, 356-359.
9. BRAY, D. W. The prediction of behavior from two attitude scales. *J. abnorm. soc. Psychol.*, 1950, 45, 64-84.

10. CAMPBELL, D. T. The indirect assessment of social attitudes. *Psychol. Bull.*, 1950, 47, 15-38.
11. CRONBACH, L. J. Further evidence on response sets and test designs. *Educ. psychol. Measmt.*, 1950, 10, 3-31.
12. CRONBACH, L. J., & GLESER, G. Assessing similarity between profiles. *Psychol. Bull.*, 1953, 50, 456-473.
13. ELINSON, J. Attitudinal intensity in relation to personality status in a military situation. Unpublished doctoral dissertation, George Washington Univer., 1954.
14. FREEMAN, R. V., & GRAYSON, H. M. Maternal attitudes in schizophrenia. *J. abnorm. soc. Psychol.*, 1955, 50, 45-52.
15. GOLDSTEIN, A. P., & CARR, A. C. The attitudes of the mothers of male catatonic and paranoid schizophrenics toward child behavior. *J. consult. Psychol.*, 1956, 20, 190.
16. GORDON, L. V. Some interrelationships among personality item characteristics. *Educ. psychol. Measmt.*, 1953, 13, 264-272.
17. GREEN, B. F. Attitude measurement. In G. Lindzey (Ed.), *Handbook of social psychology*. Cambridge, Mass.: Addison-Wesley, 1954. Pp. 335-369.
18. GUTTMAN, L., & FOA, U. G. Social contact and an intergroup attitude. *Publ. Opin. Quart.*, 1951, 15, 43-53.
19. HARRIS, D. B., GOUGH, H. G., & MARTIN, W. E. Children's ethnic attitudes: II. Relationship to parental beliefs concerning child rearing. *Child Developm.*, 1950, 21, 169-181.
20. HUBERT, M. A., & BRITTON, J. H. Attitudes and practices of mothers rearing their children from birth to the age of two years. *J. Home Econ.*, 1957, 49, 208-219.
21. JENKINS, R. L. The schizophrenic sequence: withdrawal, disorganization, psychotic reorganization. *Amer. J. Orthopsychiat.*, 1952, 22, 738-748.
22. KELLY, E. L. Consistency of the adult personality. *Amer. Psychol.*, 1955, 10, 659-681.
23. KELMAN, H. C. Attitude change as a function of response restriction. *Human Relat.*, 1953, 6, 185-214.
24. KLEBANOFF, L. B. Attitudes of mothers of schizophrenic, brain injured and retarded, and normal children. Unpublished doctoral dissertation, Boston Univer., 1957.
25. KUDER, F. G. Expected developments in interest and personality inventories. *Educ. psychol. Measmt.*, 1954, 14, 265-271.
26. LAPIERE, R. T. Attitudes vs. actions. *Soc. Forces*, 1934, 13, 230-237.
27. MCFARLAND, R. L. A replication of Mark's study of the attitudes of mothers of male schizophrenics. *J. consult. Psychol.*, in press.

CHILD DEVELOPMENT

28. McNEMAR, Q. Opinion-attitude methodology. *Psychol. Bull.*, 1946, 43, 289-374.
29. MANN, V. D. A study of the attitudes of mothers of cerebral palsied children toward child adjustment. Unpublished doctoral dissertation, American Univer., 1957.
30. MARK, J. C. The attitudes of the mothers of male schizophrenics toward child behavior. *J. abnorm. soc. Psychol.*, 1953, 48, 185-189.
31. READ, K. H. Parents' expressed attitudes and children's behavior. *J. consult. Psychol.*, 1945, 9, 95-100.
32. REMMERS, H. H. *Introduction to opinion and attitude measurement*. New York: Harper, 1954.
33. RUNDQUIST, E. A. Form of statement in personality measurement. *J. educ. Psychol.*, 1940, 31, 135-147.
34. SCHAEFER, E. S., & BELL, R. Q. Development of a parental attitude research instrument. *Child Develpm.*, 1958, 29, 339-361.
35. SCHAEFER, E. S., & BELL, R. Q. Structure of attitudes toward child-rearing and the family. *J. abnorm. soc. Psychol.*, 1957, 54, 391-395.
36. SCHANCK, R. L. A study of a community and its groups and institutions conceived of as behaviors of individuals. *Psychol. Monogr.*, 1932, 43, 1-133.
37. SHOBEN, J. R., JR. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, 39, 101-148.
38. STODGILL, R. M. The measurement of attitudes toward parental control and the social adjustment of children. *J. appl. Psychol.*, 1936, 20, 359-367.
39. STOTT, L. H., & BERSON, M. P. Some changes in attitudes resulting from a preparental education program. *J. soc. Psychol.*, 1951, 34, 191-202.
40. STOUFFER, S. A., et al. *Measurement and prediction*. Princeton, N. J.: Princeton Univer. Press, 1950.
41. SYMONDS, P. M. *The psychology of parent-child relationships*. New York: Appleton-Century, 1939.
42. THOMAS, R. C. *Mother-daughter relationships and social behavior*. Washington, D. C.: Catholic Univer. of America Press, 1955.
43. TIETZE, T. A study of mothers of schizophrenic patients. *Psychiatry*, 1949, 12, 55-65.
44. WILEY, J. H. A scale to measure parental attitudes toward certain aspects of children's behavior. Unpublished doctoral dissertation, Univer. of Calif., 1950.

DEVELOPMENT OF A PARENTAL ATTITUDE RESEARCH INSTRUMENT¹

EARL S. SCHAEFER *and* RICHARD Q. BELL
*National Institute of Mental Health*²

This paper will present (a) the thesis that parental attitudes toward child-rearing and the family life are an important influence in the personality development of the child, (b) a review of objective studies of parental attitudes toward child rearing, (c) a rationale for the development of homogeneous measures of concepts concerning parental attitudes and a description of the process of concept development, (d) the results of an analysis of item and scale efficiency, and (e) correlations of the attitude measures with education.

The current extensive interest in the influence of the parent-child relationship upon the personality development of the child arose primarily through psychotherapeutic work with adults and the clinical studies of families by child guidance clinics. Many of the earlier studies have been reviewed by Symonds (37) and Radke (31). Orlansky's review (29) of literature relevant to the hypothesis that types of infant care determine personality structure suggested the need for an emphasis upon the total pattern of experiences of children rather than upon specific infant care practices in the study of environmental influences upon personality development. Symonds (37) also concluded that the quality of the parent-child relationship is of crucial importance. Mead's theory of the genesis of the self is a concise statement of this point of view: "The individual experiences himself as such, not directly, but only indirectly, from the particular standpoints of other individual members of the same social group, or from

¹ The writers wish to express deep appreciation to Comdr. Thomas E. Cone, USN, Chief of Pediatrics, U. S. Naval Medical Center, Bethesda, Maryland, for his cooperation and assistance in making this study possible.

² National Institutes of Health, Public Health Service, U. S. Department of Health, Education, and Welfare.

CHILD DEVELOPMENT

the generalized standpoint of the social group as a whole to which he belongs" (26, p. 138). Since the most extensive and intensive social interactions of the child during crucial developmental stages occur within the family and especially with the mother, the mother-child relationship would be of major importance in personality development.

Recent trends in personality theory provide a rationale for the conceptualization and quantification of maternal attitudes. Bronfenbrenner (6) has noted a convergence upon a set of related concepts which imply a tendency to orient and act with respect to objects. In action theory these tendencies have been called need-dispositions (30) while similar concepts have variously been called dispositions (1), dispositional tendencies (17), needs (28), belief-value matrices (38), and dynamic systems (18). These concepts include emotional, cognitive, and motivational elements and imply *stable* genotypical personality characteristics. Let us assume that such need-dispositions are related to the type of relationship which a woman develops with her child. Development of measures of those components of her personality which are relevant to her role as a mother would permit prediction of her behavior with her child and the future personality adjustment of the child. The hypothesis that such need-dispositions can to some degree be determined by measuring the attitudes of parents toward child-rearing and the family is the motivation for this study.

The importance of maternal attitudes to the development of the child has been generally accepted. Bettelheim (3) has pointed out that if parents who rigidly attempt to impose accelerated developmental goals upon children are told that early toilet-training is undesirable they may delay such training but may continue to accelerate the child in talking, reading, and other important areas of development. This example illustrates that an underlying attitude may influence a great variety of parent behaviors. Many researchers have implicitly assumed that objective attitude measures could give valid information on such parental attitudes. The potential usefulness of such objective measures justifies their development. When such measures are available, their construct validity can be objectively determined.

If objective attitude measures can be demonstrated to be indices of certain patterns of maternal behavior, attitude measurement would be a valuable supplement to interview and observation as methods for the study of maternal behavior. The Maccobys' discussion of the relative merits of interviews and questionnaires is relevant to the question of validity; they state: "It has been widely assumed that the interview is superior to the questionnaire in many ways, and should be used where resources permit. Certainly the interview *must* be used at exploratory stages; for the later stages of research, the meager evidence so far available does not support the superiority of the interview. . . . It is entirely possible that the interview is superior for some subject matters, the questionnaire for others" (21, p. 483). If methods for coding interviews were well developed, ample time

for thorough interviews were available, and well trained interviewers could be obtained, it seems likely that interviews would give more valid information. However, the speed, efficiency and objectivity of the questionnaire make it a desirable method at this stage of research on parent-child relationships.

REVIEW OF PREVIOUS STUDIES

Turning to the empirical evidence for the validity of this approach to the study of parent-child relationships, we find a number of researches which have studied relationships between parental attitudes and the personality adjustment of children. Stogdill (36) was among the first to attempt the objective measurement of parental attitudes with two questionnaires which measured attitudes toward parental control and attitudes toward social behavior of children. Read (32) studied the relationship of these variables with a behavior inventory for nursery school children by comparing the responses of 32 parents who were tested with these measures with the rated behavior of their 20 children. Although she concluded that liberalism in views on parental control is related to child behavior, correlations between scores on parental attitude scales and summarized ratings were not significant when computed from tabular data presented in the research report. Anderson (2) reported a study completed by Miles (27) in which it was found that parents' attitudes are related to the leadership status and social acceptance of the child by the school group. He reported that subscales measuring overprotection, dominance by the parent, and encouragement of social development are related to the child's social adjustment. The effect of a significant difference in education between parent groups was not partialled out. Shoben (35) administered an inventory of attitudes toward child-rearing to 50 mothers of children in mental hygiene clinics and juvenile courts and 50 mothers of normal children. Significant correlations between three parental attitude scales and children's adjustment were shown on cross-validation with comparable groups. Scales measuring maternal dominance, possessiveness, and ignoring were approximately equal in power of differentiation. Partial correlations computed from data available in the published report indicate that a significant difference in educational level between parents of problem and nonproblem children did not materially affect the validity coefficients obtained.

Radke (31) constructed scales of parental attitudes for the areas of autocratic or democratic control of the child, amount and area of restrictions on the child, amount of freedom, severity or mildness of punishment, rapport between parent and child, the relative responsibility of the parents for the child's discipline, and sibling compatibility or incompatibility. No reliabilities for these scales were given but the intercorrelations of the measures were low. Radke stated that relationships between these home atmosphere scales and child behavior ratings by preschool teachers sug-

gested that unfavorable conduct of children is related to autocratic, restrictive, and severe discipline. These conclusions were based on finding 11 differences significant at or beyond the .05 level, where seven would be expected by chance, in 140 comparisons of 20 different behavior ratings of preschool children and seven attitude scales administered to their parents. The extreme homogeneity of the group of parents on education and other characteristics may have operated to depress relationships.

Mark (25) assembled items which seemed relevant to the concept of the schizophrenogenic mother and administered them to 100 mothers of hospitalized schizophrenics and 100 mothers of controls who were matched for age, religion, education, socioeconomic status, and age of their sons. He found that 67 of the 139 items differentiated the groups beyond the .10 level of confidence. A replication of this study by McFarland (23) failed to find significant differences between mothers of normals and of schizophrenics. Computations based on data available from both the Mark and McFarland studies indicate that the number of items differentiating the two populations of mothers of schizophrenics was significantly greater than expected on a chance basis, while the differences between the mothers of the nonschizophrenic individuals in both studies were close to the number expected on a chance basis. Freeman and Grayson (10) measured the attitudes of 50 mothers of schizophrenics and 50 mothers of normals. They report that the possessive and unclassified groups of items from the Shoben inventory (35) differentiated their groups at the .05 level of confidence although the dominating and ignoring subscales failed to reach that level. Identifying data relative to age, religion, education, or socioeconomic level were not obtained for mothers in the control group who were contacted by volunteer students, nurses and aides; hence, it is not possible to evaluate the extent to which educational or other differences contributed to the differences obtained. Harris, Gough, and Martin (13) also tested the relationship between a priori scales which measured authoritarianism, permissiveness, parent-child integration, parental rigidity, and good judgment (the responses judged by the writers to be most desirable) and an 18-item attitude scale toward Negroes which was given to the children of these parents. The number of items differentiating high and low prejudice groups differed only slightly from what would be expected by chance at the significance levels cited. Empirical parental attitude scales constructed from the data were not cross-validated.

Shapiro (34) conducted a study of correlates of parental attitudes on child-rearing with other attitudes and objective measures on parents. He found that attitude items which approve the free expression of a child's desires are related to objective attitude measures of "Radicalism" with parents who are less restrictive being more radical in their political opinions. Items which he designated as "Analytical" dealing with anal, oral, and sibling-rivalry situations were less related to these attitudes. He concluded that attitudes toward child-rearing are related to other social attitudes and

that different types of parental attitude items are differentially related to other personality measures of the parent. Block (4), in a more comprehensive study of restrictive attitudes toward child-rearing, found a scale of paternal restrictiveness significantly correlated with 122 of 498 other tests and measures which were included in a personality assessment of 100 military officers. Paternal restrictiveness was also related to Q-sort personality evaluations which were made by eight staff psychologists who had observed the subjects during the procedures of the assessment program. A comparison of the 20 most restrictive and the 20 most permissive subjects revealed that "Restrictive fathers tend to be constricted, submissive, suggestible individuals with little self-assurance. Permissive fathers were evaluated as self-reliant, ascendant, and functioning effectively. . . ." Restrictive attitudes were also positively related to measures of ethnocentrism and fascism.

Goldstein and Carr (11), using some items previously used by Mark, administered a 56-item questionnaire to 34 mothers of adult males diagnosed as catatonic and 26 mothers of adult males diagnosed as paranoid and found that the number of differentiating items did not exceed chance expectations. Mothers of catatonics differed significantly in reports of inability to complete the questionnaire and in the number of answers omitted.

Croty (8) failed to find any correlations in excess of .25 between measurements of achievement motivation, obtained from a psychological insight questionnaire administered to 95 male high school students, and attitude measures obtained from their mothers. The attitude questionnaire, Final Form IV of the Parental Attitude Research Instrument (PARI), which is reported in this paper, was mailed to the mothers and a return of 78 per cent was obtained.

Klebanoff (16) administered Form IV of PARI to 15 mothers of schizophrenic children, 15 mothers of retarded or brain injured children, and 26 mothers of normal children. Significant differences in scales related to an overpossessiveness factor appeared between mothers of normals and mothers of schizophrenics as well as between mothers of normals and mothers of retarded or brain injured children. Differences remained after statistical control of education differences between groups. No differences could be demonstrated between mothers of schizophrenic children and mothers of mentally retarded or brain damaged children on factor scores, but differences between groups on individual scales were not tested.

Mann (24) also administered Final Form IV of PARI as well as questionnaires developed by Shoben (35) and Mark (25) to mothers of children with cerebral palsy and closely matched mothers of normal children. Mothers of normals scored significantly higher on the Ignoring scale from Shoben, but differences on the Possessive and Dominant scales were not significant. Mothers of children with cerebral palsy scored significantly higher (showed more agreement) on the PARI scales Seclusiveness, Strictness, Intrusiveness, Acceleration, Encouraging Verbalization, Equalitarian-

ism, and Comradeship and Sharing. Mothers of normals scored significantly higher on Marital Conflict, Irritability, Rejection of the Homemaking Role, and Avoidance of Communication. Eleven of the 23 PARI scales yielded differences significant at or beyond the .05 level.

This review of studies of parental attitudes toward child-rearing reveals great interest in the hypothesis that objectively measured attitudes toward child-rearing are significantly related to personality development of children. Most of the studies quoted here are interpreted by their authors as supporting the hypothesis. Despite this consensus the research findings do not seem to be conclusive due to problems of matching experimental and control groups (10, 35), the use of ad hoc scales without adequate test development (31), the attribution of significance to a small number of items out of a larger group under circumstances where such items might well have reflected chance relationships (10), the marginal significance of the findings of several studies (10, 13, 31, 32), as well as the general absence of cross-validation of findings of marginal significance.

Several of the studies reviewed, as well as a study by Wiley (40), have developed measures of attitude concepts such as parental dominance, control, or restrictiveness; possessiveness; ignoring; and parent-child rapport or integration. The conceptual schemes have often been too limited, concept definition has often been too vague, or adequate test development has been neglected. The review of the literature suggested a need for the development of a comprehensive conceptualization and quantification of this domain which would provide a research instrument for future studies.

CONCEPTUAL ANALYSIS OF PARENTAL ATTITUDES

Both clinical studies, in which various syndromes of parental attitudes have been related to differential personality development of children, and the statistical theory of differentiation and prediction suggest the need for development of measures of specific attitude concepts composed of homogeneous groups of items. Theoretically, such an inventory would be more useful than a test composed of individual items or one which would give only one score of pathogenicity of parental attitude. This opinion is supported by Holzinger's statement "... a single average as a complete summarization is justified only if the data are of rank one, that is, if only one common factor is involved" (14). Cronbach (7) also suggests that it is often advisable to obtain less reliable measures of a number of relevant variables containing specific variance rather than to investigate a limited number of very reliably measured variables. Brogden (5) has demonstrated that differentiation between groups can be significantly increased by using several predictors which are independently related to a criterion even if the predictors correlate as much as .80 with one another. Guttman's (12) criticisms of heterogeneous composite scores also suggest that the use of

only one composite score in the area of parental attitudes would tend to reduce differentiating power of a parental attitude research instrument and would obscure interpretation of results.

With these considerations in mind a conceptual analysis of the domain of parental attitudes was carried out as a basis for developing measures of attitudes. The relatively heavy saturation of statistically significant items in the studies by Mark (25) and Shoben (35) prompted sorting of their item pools by three clinical psychologists into subscales or groups of items which seemed psychologically homogeneous. In conference these conceptual analyses were discussed and a group decision was made to use the concepts given in Table 1. This conceptual scheme was used to classify all items of the two studies into one of the categories given. This analysis, although based upon varying numbers of items in the various categories, strongly suggested that some categories have higher discriminating power than others.

It also became evident, in the process of sorting, that items which make what seem to be either purely factual statements or which state attitudes

TABLE 1

CONCEPTUAL ANALYSIS OF ITEMS FROM MARK AND SHOBN STUDIES

	NO. OF ITEMS DIFFERENTIATING BEYOND .05 LEVEL OF SIGNIFICANCE		NO. OF ITEMS NOT DIFFERENTIATING		PER CENT OF ITEMS DIFFER- ENTIATING
	Mark's Study	Shoben's Study	Mark's Study	Shoben's Study	Combined
<i>Parental Attitudes</i>					
Strictness	3	6	0	0	100
Intrusiveness	3	2	0	0	100
Martyrdom	2	1	0	0	100
Breaking the Will	2	5	1	0	88
Infantilization	6	5	2	0	87
Deification	6	10	4	0	80
Suppression of Aggression ...	1	2	1	0	75
Child-dominated Home	1	9	0	4	71
Nonpunishment	3	2	2	0	71
Suppression of Sexuality	6	4	1	3	71
Ignoring the Child	4	5	4	2	60
Avoidance of Tenderness	3	0	1	1	60
Absolute Authority of Parent .	2	13	7	5	56
Social Isolation of Child	2	6	5	3	50
Acceleration	4	1	6	0	45
Statement of Fact	0	0	8	6	00
Expression of Positive Attitude	0	0	4	14	00
Unclassified Items	19	14	26	25	39

that would be generally endorsed by psychologists as expressing approved attitudes toward child-rearing do not differentiate effectively between parents of normals and parents of maladjusted children. This is in keeping with the report of Stogdill (36) who noted that often parents will agree with apparently contradictory attitudes if one statement expresses what appears to be a healthy attitude while the other states its converse. Harris, Gough, and Martin (13) also report that items stating healthy attitudes are typically poor discriminators. We found that scales which state approved attitudes toward child-rearing typically have very poor reliabilities because there is a strong tendency for all persons to agree. As a result there is little variability and poor discrimination on so-called healthy scales.

This finding led us, in constructing a set of scales for this domain, to concentrate our efforts upon the development of the more differentiating scales which state attitudes contrary to the usually approved child-rearing opinions. It was soon found that tests composed entirely of such scales left many mothers very dissatisfied. They frequently expressed the feeling that they did not see any items with which they could agree and therefore could not indicate their own ideas. To correct this we included in our own inventory a number of scales which we should expect, from previous studies, to have relatively low reliabilities and poor discrimination. These so-called "Rapport" scales are Equalitarianism, Expression of Affection, Comradeship and Sharing, Encouraging Verbalization, and Autonomy of the Child.

Our procedure in developing additional scales was as follows. After developing a concept on the basis of theory or formulating a concept from the writings of others, we attempted an operational definition by writing items designed to measure the concept. Scales which were thus developed were tried out in short exploratory forms on samples of 25 mothers. Items were eliminated if they did not reveal variation in attitudes or did not vary with other items which were used to define the scale. Promising items were noted, and new items were written based upon hypotheses about the successful items. Usually two or three successive trial forms, each of which utilized the insights derived from previous trial forms, were used in developing an item pool which was then standardized on samples of 100. This procedure permitted a process of hypothesis development and testing which maximized the amount of new information which could be used in developing an effective item pool, and did not impose statistical rigor at the level of hypothesis formation. Use of successive samples, and confirmation of small sample work with larger samples in the final forms, supported the hypothesis that this process resulted in improvement in operational definition of the concepts.

The results of this sequential experimentation compared with results which would have been obtained by a simple elimination of items from an original item pool is shown in the development of the scale Harsh Pun-

ishment. The reliability or internal consistency of the original pool of 10 items for the scale Harsh Punishment was .12, which could have been improved to the level of .15 by elimination of four items. However, by writing new items to follow up promising leads this scale was finally developed into an eight-item form with a reliability of .82. Improvements in scale reliability were frequently rather dramatic following detection of the psychological basis of the success of discriminating items.

The selection of concepts for development was based upon a search of the literature for hypotheses as to which areas would give additional information relevant to parent-child relationships. As a result of the hypothesis that information on marital adjustment would assist greatly in understanding the home environment of the child, several scales were developed which measured relevant concepts. The scale Marital Conflict contains items which were intended to reveal the existence of tensions indirectly through obtaining the respondents' judgments about how much quarreling and dissension might be expected in marriages in general. The scale Rejection of the Homemaking Role contains items which state the unhappiness of a woman at being shut up in a home, and her dissatisfaction with the duties of caring for the home and children. The scale Irritability was developed around items which indicate that children "get on a woman's nerves" and that any woman would "blow her top" frequently in the difficult job of managing a home. This scale may be related to several sources of tension but it seems feasible that it would also reflect tensions in the marital relationship. The scale Inconsiderateness of the Husband contains items which state how inconsiderate husbands in general are. It is intended to permit projection of feeling about the spouse. Ascendancy of the Mother, which measures the extent to which the woman plans to dominate the family, was also developed to investigate a possible source of marital conflict.

The two scales, Encouraging Verbalization and Avoidance of Communication, were written to determine whether the parent would permit or encourage the child to talk about his anxieties, conflicts, hostilities, and disagreements with parental policies. Ignoring the Baby was designed to determine how much affection and attention the mother thought was appropriate for the care of the infant. These scales which measure attitudes toward communication and toward care of the infant hopefully should measure the warmth and permissiveness of the parent-child relationship. The scale Fear of Harming the Baby was developed out of interest in clinical reports of the eruption of unconscious hostile feelings toward their infants in disturbed mothers. Seclusion of the Mother was designed to measure the social adjustment of the mother, since it would presumably influence the socialization of the child. Dependency of the Mother is an attempt to determine the extent to which the mother looks to others for security and support. Fostering Dependency was designed to measure another aspect of overprotection or overpossessiveness. Harsh Punishment

CHILD DEVELOPMENT

attempts to determine the degree to which the mother feels punishment is an effective method of influencing and controlling children. Excluding Outside Influences, which is designed to measure family ethnocentrism, may be related to parental control and authoritarian attitudes.

Abdication of the Parental Role reflects an attitude that parents can do relatively little with children and that most parental efforts are wasted. This scale, with Rejection of the Homemaking Role and Irritability, was included in an effort to determine the extent to which the mother rejects or ignores her children. We felt it important to explore this area since the items utilized in previous studies seemed to be more relevant to clinical descriptions of parents of withdrawn or schizophrenic children than to parents of aggressive or delinquent children. Demand for Activity items measure the extent to which the parent feels children should be kept busy and rushing at all times.

Although no systematic effort was made to develop generalizations concerning successful item-writing, it was frequently found that vague items which presumably enhance projective responses could be used to explore difficult areas. Rationalizations used by parents of disturbed children were used with the hope that psychopathogenic parents would agree while the critical judgment of more normal parents would cause them to reject the statement. An example is the item, "Children need some of the natural meanness taken out of them," in the scale Breaking the Will. The apparent truism seems to give the respondent with a latent but defended attitude a feeling that it may be socially acceptable to agree. Cliches, colloquialisms, and conventional affect-laden phrases were often useful. Examples of these are phrases in the scale Irritability such as "will get on any woman's nerves," "can't stand their children a moment longer," "it's natural for a woman to blow her top." An example of attaching the aura of self-righteousness to an item is, "A wise parent will teach a child early just who is boss." Such techniques were used in item writing with the hypothesis that threshold for item acceptance would be changed differentially in normal and deviant parents. No empirical data are available, however, to test this hypothesis.

THE SAMPLES

During two different two-month periods women were tested within one to four days following delivery of their babies at a military hospital in the metropolitan area of Washington, D.C. Mothers who were not sufficiently recovered from sedation at the time of test administration, mothers with less than an eighth grade education, and non-English speaking mothers were excluded. Age and educational levels of these groups are reported in Tables 2 and 3. These data are reported as a partial assessment of the representativeness of these groups. Test instructions were given on

the ward, each mother working individually while remaining in her bed. Trial Form III was given to 100 freshman and sophomore student nurses. Final Form IV, which was used to determine the reliability of the five item scales, was given to a group similar to the samples for Trial Forms I and II.

TABLE 2
EDUCATION OF PARI SAMPLES CONTRASTED WITH 1950 CENSUS DATA*

	PRIMIPARAE N = 200 Per Cent	MULTIPARAE N = 200 Per Cent	1950 CENSUS Per Cent
Grammar School	2	2	13
Some High School	10	13	25
High School Graduate	50	49	42
Some College	23	24	12
College Graduate	13	10	6
No information	2	2	2

* 1950 census data on education of all white females, urban and rural, within the age range of 18 to 40 years.

TABLE 3
AGES OF PARI SAMPLES CONTRASTED WITH 1952 VITAL STATISTICS*

P A R I S A M P L E S			1 9 5 2 V I T A L S T A T I S T I C S		
AGE Years	PRIMIPARAE N = 200 Per Cent	MULTIPARAE N = 200 Per Cent	AGE Years	PRIMIPARAE Per Cent	MULTIPARAE Per Cent
-18 ..	9	2	-19 ..	25	4
19-21 ..	36	12	20-24 ..	44	26
22-24 ..	34	25	25-29 ..	20	33
25-27 ..	9	15	30-34 ..	8	23
28-30 ..	5	15	35- ..	3	14
31-33 ..	2	13			
34-36 ..	3	10			
37- ..	0	7			
No information	2	1			

* Class intervals cannot be compared directly since PARI data were obtained in class intervals which differed from vital statistics class intervals.

CHILD DEVELOPMENT

TABLE 4

RELIABILITY ESTIMATES FOR VARIOUS PARI SAMPLES

Scale	Group:	PRIMIPARAE			MULTIPARAE			NURSES	
	Form:	Trial I	Trial II	Final IV	Trial I	Trial II	Final IV	Trial III	Trial III
	N:	100	100	60	100	100	60	100	60
	Formula:	KR20	KR20	KR20*	KR20	KR20	KR20	KR20	PPM†
	No. items:	10	8	5	10	8	5	8	8
<i>Equalitarianism</i>		.37		.40	.58		.46		
Nonpunishment		.67			.54				
Avoidance of Tenderness		.62			.58				
Suppression of Aggression		.77		.67	.80		.67	.54	.48
Breaking the Will		.69		.60	.73		.68	.63	.74
Autonomy of the Child		.66			.62			.61	.18
Ignoring the Child		.48			.43				
Strictness		.78		.69	.79		.74	.74	.76
Intrusiveness		.86		.76	.85		.73	.84	.63
Expression of Affection		.49			.57				
Infantilization		.63			.70			.57	.47
Suppression of Sex		.66		.63	.79		.67	.58	.51
Acceleration of Development		.71		.65	.61		.70	.64	.51
Comradeship and Sharing		.62		.60	.55		.40	.62	.62
Deification		.77		.55	.69		.77	.79	.75
Martyrdom		.78		.67	.85		.67	.65	.73
Encouraging Verbalization			.62	.34		.67	.45	.61	.44
Seclusion of the Mother		.77		.70		.75	.70	.61	.68
Harsh Punishment		.82				.83		.76	.66
Dependency of the Mother		.69		.54		.66	.47		
Fear of Harming the Baby		.74		.72		.66	.70		
Fostering Dependency		.75		.77		.76	.73	.60	.71
Marital Conflict		.71		.69		.68	.60	.64	.64
Irritability		.77		.63		.77	.54	.73	.79
Excluding Outside Influences		.77		.63		.77	.74	.67	.72
Rejection of the Homemaking Role		.79		.68		.73	.68	.72	.62
Avoidance of Communication		.74		.57		.71	.64	.59	.53
Ignoring the Baby		.70				.63		.77	.66
Ascendancy of the Mother		.86		.71		.77	.76	.74	.60
Inconsiderateness of the Husband		.81		.69		.83	.73		
Approval of Activity		.72		.45		.71	.64	.66	.64
Abdication of the Parental Role		.66				.43		.64	.78

* Cross-validated reliability of five best items selected from Form I or II calculated with Kuder-Richardson Formula 20.

† Pearson product-moment test-retest reliability.

RELIABILITIES

The results of investigations of internal consistency and test-retest reliabilities of the scales are reported in Table 4. Reliabilities were estimated with Kuder-Richardson Formula 20 (19) for the scales included in Trial Forms I and II for both multiparae and primiparae. In an attempt to develop a more efficient test, the five most reliable items for each scale were selected for Final Form IV which consists of 23 five-item scales. Internal consistency reliability coefficients were calculated for these five-item scales on new samples of 60 multiparae and 60 primiparae. In general, the reliabilities were again satisfactory for multivariate research on group differences. Scales from both Trial Forms I and II were included in Trial Form III which was administered to 100 student nurses who were highly homogeneous in age, education, and socioeconomic background. Despite the homogeneity of the sample, internal consistency reliability coefficients were again satisfactory. Test-retest reliabilities, which were calculated for 60 of the group who were retested after a period of three months, were generally good although a few scales on which there was very little variability in score in this group had appreciably lower test-retest reliability as compared to internal consistency reliability.

ITEM ANALYSIS

Although the reliability estimates for eight- and 10-item scales in the initial forms were satisfactory, it seems probable that individual items were not effectively measuring the concepts in which they had been included. The same data upon which reliabilities had been calculated were therefore utilized to determine the correlation of each item with the total score of the scale in which it had been included. An analysis described by Wherry and Winer (39), which estimates the loading of an item on the centroid of the scale in which it is included, was performed.³ These data permitted the selection of 23 subscales of five items each for Final Form IV of the Parental Attitude Research Instrument (PARI) which is reported in the appendix with instructions which were used in administering the inventory. In Final Form IV the items within a scale are cyclically arranged so that the first subscale includes items 1, 24, 47, 70, and 90, the second subscale items 2, 25, 48, 71, and 94, etc. A more detailed description of the test and its use are reported in an unpublished manuscript (33).

³ A table containing the total set of items together with their loadings on the scale centroids has been deposited as Document number 5559 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. A copy may be secured by citing the Document number and by remitting \$3.75 for photocopies, or \$2.00 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

CHILD DEVELOPMENT

TABLE 5

PARI SCALE CORRELATIONS WITH EDUCATIONAL LEVEL

<i>Attitude Scale</i>	<i>Primiparae</i> (<i>N</i> = 100)	<i>Multiparae</i> (<i>N</i> = 100)
Equalitarianism	-.11	-.02
Nonpunishment	-.03*	-.32
Avoidance of Tenderness	-.09	-.20
Suppression of Aggression	-.33	-.34
Breaking the Will	-.20	-.42
Autonomy	-.31	-.12
Ignoring the Child	-.06	-.15
Strictness	-.13	-.24
Intrusiveness	-.30	-.40
Expressing Love and Affection	-.02	-.06
Infantilization	-.29	-.42
Suppression of Sexuality	-.33	-.48
Acceleration of Development	-.31	-.46
Comradeship and Sharing	-.05	.14
Deification of the Parent	-.26	-.52
Martyrdom	-.38	-.42
Encouraging Verbalization	.09	.05
	(<i>N</i> = 96)	(<i>N</i> = 97)
Seclusiveness of the Mother	-.30	-.22
Harsh Punishment	-.29	-.27
Dependency of the Mother	-.28	-.11
Fear of Harming the Baby	-.27	-.27
Fostering Dependency	-.33	-.31
Marital Conflict	-.35	-.17
Irritability	-.22	-.26
Excluding Outside Influences	-.25	-.15
Rejection of the Homemaking Role	-.07	-.18
Avoidance of Communication	-.30	-.20
Ignoring the Baby	-.11	-.03
Ascendancy of the Mother	-.21	-.19
Inconsiderateness of the Husband	-.20	-.20
Approval of Activity	-.08	-.24
Abdication of the Parental Role	-.33	-.20

* Positive correlation should be interpreted as meaning a person with a high education level agrees with the attitude expressed by the scale.

Although predictive validity studies have not yet been done on the PARI, the studies surveyed in the review of literature could be cited as evidence supporting the concurrent validity of this general approach to the study of parent-child relationships. The content validity or the adequacy of the concepts of parental attitudes toward child-rearing can be determined by attempting to add concepts which would add new variance to measurement of this domain. Adequate evidence of the construct validity of these scales must be based upon research on various correlates of these measures.

CORRELATIONS OF SCALES WITH EDUCATIONAL LEVEL

Previous investigations (9, 20) have found that differences in attitudes toward child-rearing are related to the social class level of the respondents. A recent investigation of the interrelationships of various indices of social class level led to the conclusion that level of educational achievement is highly related to other indicators of socioeconomic level (15). These attitude scores were therefore correlated with self-reports of number of years of school completed successfully. The findings, which are reported in Table 5, seem to be congruent with the results reported by Maccoby *et al.* (20) in that mothers of the lower educational levels tend to have less approved attitudes toward children. The correlations of education with attitude for the scales which state disapproved attitudes toward child-rearing range from $-.03$ to $-.38$ for primiparae with a mean of $-.24$, and from $-.03$ to $-.52$ for multiparae with a mean of $-.27$. Although many of these correlations are statistically significant at beyond the .05 level, the absolute magnitude of the correlations does not justify the interpretation of the variations in attitude as merely reflecting differences in educational level.

SUMMARY

Theories of parental influence upon development of children and a review of previous research on the relationship of parental attitudes to the personality adjustment of children suggested the need for the development of a set of homogeneous measures of parental attitudes. A set of 32 concepts were selected which were derived from previous studies and from a search of the literature on parent-child relationships.

Attitude scales of five to 10 items which gave satisfactory reliability for research purposes were developed with an iterative technique of attitude measurement. Many of these measures were found to be related to education, with mothers of higher education having more usually approved attitudes toward child-rearing. It is suggested that logically, psychologically, and empirically homogeneous scales of attitude toward child-rearing will be useful in investigating theories of the influence of maternal attitudes upon development of the child.

CHILD DEVELOPMENT

APPENDIX A

SAMPLE COPY OF FINAL FORM IV (PARI)*

INVENTORY OF ATTITUDES ON FAMILY LIFE AND CHILDREN

Read each of the statements below and then rate them as follows:

A strongly agree	a mildly agree	d mildly disagree	D strongly disagree
------------------------	----------------------	-------------------------	---------------------------

Indicate your opinion by drawing a circle around the "A" if you strongly agree, around the "a" if you mildly agree, around the "d" if you mildly disagree, and around the "D" if you strongly disagree.

There are no right or wrong answers, so answer according to your own opinion. It is very important to the study that all questions be answered. Many of the statements will seem alike but all are necessary to show slight differences of opinion.

- | | A | a | d | D |
|---|---|---|---|---|
| 1. Children should be allowed to disagree with their parents if they feel their own ideas are better. | A | a | d | D |
| 2. A good mother should shelter her child from life's little difficulties. | A | a | d | D |
| 3. The home is the only thing that matters to a good mother. | A | a | d | D |
| 4. Some children are just so bad they must be taught to fear adults for their own good. | A | a | d | D |
| 5. Children should realize how much parents have to give up for them. | A | a | d | D |
| 6. You must always keep tight hold of baby during his bath for in a careless moment he might slip. | A | a | d | D |
| 7. People who think they can get along in marriage without arguments just don't know the facts. | A | a | d | D |
| 8. A child will be grateful later on for strict training. | A | a | d | D |
| 9. Children will get on any woman's nerves if she has to be with them all day. | A | a | d | D |
| 10. It's best for the child if he never gets started wondering whether his mother's views are right. | A | a | d | D |
| 11. More parents should teach their children to have unquestioning loyalty to them. | A | a | d | D |
| 12. A child should be taught to avoid fighting no matter what happens. | A | a | d | D |
| 13. One of the worst things about taking care of a home is a woman feels that she can't get out. | A | a | d | D |

* Permission is granted to anyone to reproduce this material with proper acknowledgment, without permission of the authors or of the Society for Research in Child Development.

EARL S. SCHAEFER and RICHARD Q. BELL

- | | |
|--|---------|
| 14. Parents should adjust to the children some rather than always expecting the children to adjust to the parents. | A a d D |
| 15. There are so many things a child has to learn in life there is no excuse for him sitting around with time on his hands. | A a d D |
| 16. If you let children talk about their troubles they end up complaining even more. | A a d D |
| 17. Mothers would do their job better with the children if fathers were more kind. | A a d D |
| 18. A young child should be protected from hearing about sex. | A a d D |
| 19. If a mother doesn't go ahead and make rules for the home the children and husband will get into troubles they don't need to. | A a d D |
| 20. A mother should make it her business to know everything her children are thinking. | A a d D |
| 21. Children would be happier and better behaved if parents would show an interest in their affairs. | A a d D |
| 22. Most children are toilet trained by 15 months of age. | A a d D |
| 23. There is nothing worse for a young mother than being alone while going through her first experience with a baby. | A a d D |
| 24. Children should be encouraged to tell their parents about it whenever they feel family rules are unreasonable. | A a d D |
| 25. A mother should do her best to avoid any disappointment for her child. | A a d D |
| 26. The women who want lots of parties seldom make good mothers. | A a d D |
| 27. It is frequently necessary to drive the mischief out of a child before he will behave. | A a d D |
| 28. A mother must expect to give up her own happiness for that of her child. | A a d D |
| 29. All young mothers are afraid of their awkwardness in handling and holding the baby. | A a d D |
| 30. Sometimes it's necessary for a wife to tell off her husband in order to get her rights. | A a d D |
| 31. Strict discipline develops a fine strong character. | A a d D |
| 32. Mothers very often feel that they can't stand their children a moment longer. | A a d D |
| 33. A parent should never be made to look wrong in a child's eyes. | A a d D |
| 34. The child should be taught to revere his parents above all other grown-ups. | A a d D |
| 35. A child should be taught to always come to his parents or teachers rather than fight when he is in trouble. | A a d D |
| 36. Having to be with the children all the time gives a woman the feeling her wings have been clipped. | A a d D |
| 37. Parents must earn the respect of their children by the way they act. | A a d D |
| 38. Children who don't try hard for success will feel they have missed out on things later on. | A a d D |

CHILD DEVELOPMENT

- | | |
|--|---------|
| 39. Parents who start a child talking about his worries don't realize that sometimes it's better to just leave well enough alone. | A a d D |
| 40. Husbands could do their part if they were less selfish. | A a d D |
| 41. It is very important that young boys and girls not be allowed to see each other completely undressed. | A a d D |
| 42. Children and husbands do better when the mother is strong enough to settle most of the problems. | A a d D |
| 43. A child should never keep a secret from his parents. | A a d D |
| 44. Laughing at children's jokes and telling children jokes makes things go more smoothly. | A a d D |
| 45. The sooner a child learns to walk the better he's trained. | A a d D |
| 46. It isn't fair that a woman has to bear just about all the burden of raising children by herself. | A a d D |
| 47. A child has a right to his own point of view and ought to be allowed to express it. | A a d D |
| 48. A child should be protected from jobs which might be too tiring or hard for him. | A a d D |
| 49. A woman has to choose between having a well run home and hobnobbing around with neighbors and friends. | A a d D |
| 50. A wise parent will teach a child early just who is boss. | A a d D |
| 51. Few women get the gratitude they deserve for all they have done for their children. | A a d D |
| 52. Mothers never stop blaming themselves if their babies are injured in accidents. | A a d D |
| 53. No matter how well a married couple love one another, there are always differences which cause irritation and lead to arguments. | A a d D |
| 54. Children who are held to firm rules grow up to be the best adults. | A a d D |
| 55. It's a rare mother who can be sweet and even tempered with her children all day. | A a d D |
| 56. Children should never learn things outside the home which make them doubt their parents' ideas. | A a d D |
| 57. A child soon learns that there is no greater wisdom than that of his parents. | A a d D |
| 58. There is no good excuse for a child hitting another child. | A a d D |
| 59. Most young mothers are bothered more by the feeling of being shut up in the home than by anything else. | A a d D |
| 60. Children are too often asked to do all the compromising and adjustment and that is not fair. | A a d D |
| 61. Parents should teach their children that the way to get ahead is to keep busy and not waste time. | A a d D |
| 62. Children pester you with all their little upsets if you aren't careful from the first. | A a d D |
| 63. When a mother doesn't do a good job with children it's probably because the father doesn't do his part around the home. | A a d D |

- | | |
|--|---------|
| 64. Children who take part in sex play become sex criminals when they grow up. | A a d D |
| 65. A mother has to do the planning because she is the one who knows what's going on in the home. | A a d D |
| 66. An alert parent should try to learn all her child's thoughts. | A a d D |
| 67. Parents who are interested in hearing about their children's parties, dates and fun help them grow up right. | A a d D |
| 68. The earlier a child is weaned from its emotional ties to its parents the better it will handle its own problems. | A a d D |
| 69. A wise woman will do anything to avoid being by herself before and after a new baby. | A a d D |
| 70. A child's ideas should be seriously considered in making family decisions. | A a d D |
| 71. Parents should know better than to allow their children to be exposed to difficult situations. | A a d D |
| 72. Too many women forget that a mother's place is in the home. | A a d D |
| 73. Children need some of the natural meanness taken out of them. | A a d D |
| 74. Children should be more considerate of their mothers since their mothers suffer so much for them. | A a d D |
| 75. Most mothers are fearful that they may hurt their babies in handling them. | A a d D |
| 76. There are some things which just can't be settled by a mild discussion. | A a d D |
| 77. Most children should have more discipline than they get. | A a d D |
| 78. Raising children is a nerve-wracking job. | A a d D |
| 79. The child should not question the thinking of his parents. | A a d D |
| 80. Parents deserve the highest esteem and regard of their children. | A a d D |
| 81. Children should not be encouraged to box or wrestle because it often leads to trouble or injury. | A a d D |
| 82. One of the bad things about raising children is that you aren't free enough of the time to do just as you like. | A a d D |
| 83. As much as is reasonable a parent should try to treat a child as an equal. | A a d D |
| 84. A child who is "on the go" all the time will most likely be happy. | A a d D |
| 85. If a child has upset feelings it is best to leave him alone and not make it look serious. | A a d D |
| 86. If mothers could get their wishes they would most often ask that their husband be more understanding. | A a d D |
| 87. Sex is one of the greatest problems to be contended with in children. | A a d D |
| 88. The whole family does fine if the mother puts her shoulders to the wheel and takes charge of things. | A a d D |
| 89. A mother has a right to know everything going on in her child's life because her child is part of her. | A a d D |

CHILD DEVELOPMENT

90. If parents would have fun with their children, the children would be more apt to take their advice. A a d D
91. A mother should make an effort to get her child toilet trained at the earliest possible time. A a d D
92. Most women need more time than they are given to rest up in the home after going through childbirth. A a d D
93. When a child is in trouble he ought to know he won't be punished for talking about it with his parents. A a d D
94. Children should be kept away from all hard jobs which might be discouraging. A a d D
95. A good mother will find enough social life within the family. A a d D
96. It is sometimes necessary for the parents to break the child's will. A a d D
97. Mothers sacrifice almost all their own fun for their children. A a d D
98. A mother's greatest fear is that in a forgetful moment she might let something bad happen to the baby. A a d D
99. It's natural to have quarrels when two people who both have minds of their own get married. A a d D
100. Children are actually happier under strict training. A a d D
101. It's natural for a mother to "blow her top" when children are selfish and demanding. A a d D
102. There is nothing worse than letting a child hear criticisms of his mother. A a d D
103. Loyalty to parents comes before anything else. A a d D
104. Most parents prefer a quiet child to a "scrappy" one. A a d D
105. A young mother feels "held down" because there are lots of things she wants to do while she is young. A a d D
106. There is no reason parents should have their own way all the time, any more than that children should have their own way all the time. A a d D
107. The sooner a child learns that a wasted minute is lost forever the better off he will be. A a d D
108. The trouble with giving attention to children's problems is they usually just make up a lot of stories to keep you interested. A a d D
109. Few men realize that a mother needs some fun in life too. A a d D
110. There is usually something wrong with a child who asks a lot of questions about sex. A a d D
111. A married woman knows that she will have to take the lead in family matters. A a d D
112. It is a mother's duty to make sure she knows her child's innermost thoughts. A a d D
113. When you do things together, children feel close to you and can talk easier. A a d D
114. A child should be weaned away from the bottle or breast as soon as possible. A a d D
115. Taking care of a small baby is something that no woman should be expected to do all by herself. A a d D

APPENDIX B

INSTRUCTIONS TO SUBJECTS FOR PARI

"I am In cooperation with (hospital or school authorities) I am studying what (mothers) (women) think about how children should be brought up. A lot is written on this subject in various newspaper and magazine articles. Frequently these articles are not in agreement. I thought it would be a good idea to find out what (mothers) (women) themselves think. You can help in the study by passing on your own ideas. This is voluntary. If you would contribute your ideas I will leave the form with you and return for it in approximately 20 minutes. Be frank and give your own personal views regardless of what others may think. You do not need to give your name. I would, however, like you to fill out the questions on the last page for research purposes (age, education, etc.). So as not to use too much of your time we have a list of ideas which other mothers have contributed. You merely circle one of the four letters by each statement. Circle the large "A" if you strongly agree, the small "a" if you mildly agree, the small "d" if you mildly disagree, the large "D" if you strongly disagree. If you have any ideas which you feel should be included jot them down at the end. We would appreciate having them. Others who have given us their ideas say that it is best to work rapidly. Give your first reaction. If you read and reread the statements it tends to be confusing and you can't finish in the amount of time we have."

REFERENCES

1. ALLPORT, G. W. *Personality. A psychological interpretation*. New York: Henry Holt, 1937.
2. ANDERSON, J. E. Parents' attitudes on child behavior: A report of three studies. *Child Develpm.*, 1946, 17, 91-97.
3. BETTELHEIM, B. Mental health and current mores. *Amer. J. Orthopsychiat.*, 1952, 22, 76-78.
4. BLOCK, J. Personality characteristics associated with fathers' attitudes toward child-rearing. *Child Develpm.*, 1955, 26, 41-48.
5. BROGDEN, H. E. Increased efficiency of selection resulting from replacement of a single predictor with several differential predictors. *Educ. psychol. Measmt.*, 1951, 11, 173-195.
6. BRONFENBRENNER, U. Personality. In C. P. Stone & D. W. Taylor (Eds.), *Annual review of psychology*. Stanford, Calif.: Annual Reviews Inc., Vol. 4, 1953. Pp. 157-182.
7. CRONBACH, L. J. Report on a psychometric mission to clinicia. *Psychometrica*, 1954, 19, 263-270.
8. CROTY, C. J. Maternal attitudes and their relation to son's achievement motivation. Unpublished master's thesis, Catholic Univer., 1957.

CHILD DEVELOPMENT

9. DAVIS, A., & HAVIGHURST, R. J. Social class and color differences in child-rearing. *Amer. sociol. Rev.*, 1946, 11, 698-710.
10. FREEMAN, R. V., & GRAYSON, H. M. Maternal attitudes in schizophrenia. *J. abnorm. soc. Psychol.*, 1955, 50, 45-52.
11. GOLDSTEIN, A. P., & CARR, A. C. The attitudes of the mothers of male catatonic and paranoid schizophrenics toward child behavior. *J. consult. Psychol.*, 1956, 20, 190.
12. GUTTMAN, L. Relation of scalogram analysis to other techniques. In S. A. Stouffer *et al.* (Eds.), *Measurement and prediction*. Princeton, N.J.: Princeton Univer. Press, 1950. Pp. 172-212.
13. HARRIS, D. B., GOUGH, H. B., & MARTIN, W. E. Children's ethnic attitudes: II. Relationship to parental beliefs concerning child rearing. *Child Developm.*, 1950, 21, 169-181.
14. HOLZINGER, K. J. Factoring test scores and implications for the method of averages. *Psychometrika*, 1944, 9, 155-167.
15. KAHL, J. A., & DAVIS, J. A. A comparison of indexes of socioeconomic status. *Amer. sociol. Rev.*, 1955, 20, 317-325.
16. KLEBANOFF, L. B. Attitudes of mothers of schizophrenic, brain injured and retarded, and normal children. Unpublished doctoral dissertation, Boston Univer., 1957.
17. KLEIN, G. S., & KRECH, D. The problem of personality and its theory. *J. Pers.*, 1951, 20, 2-23.
18. KRECH, D. Cognition and motivation in psychological theory. In W. Dennis (Ed.), *Current trends in psychological theory*. Pittsburgh: Univer. of Pittsburgh Press, 1951. Pp. 111-139.
19. KUDER, G. F., & RICHARDSON, M. W. The theory of the estimation of test reliability. *Psychometrika*, 1937, 2, 151-160.
20. MACCOBY, ELEANOR E., GIBBS, PATRICIA K., & the staff of the Laboratory of Human Development, Harvard Univer. Methods of child-rearing in two social classes. In W. E. Martin & Celia B. Stendler (Eds.), *Readings in child development*. New York: Harcourt, Brace, 1954. Pp. 380-396.
21. MACCOBY, ELEANOR E., & MACCOBY, NATHAN. The interview: A tool of social science. In G. Lindzey (Ed.), *Handbook of social psychology*. Vol. I. Cambridge, Mass.: Addison-Wesley, 1954. Pp. 449-487.
22. McDougall, W. *An introduction to social psychology*. London: Methuen, 1908.
23. McFarland, R. L. A replication of Mark's study of the attitudes of mothers of male schizophrenics. *J. consult. Psychol.*, in press.
24. MANN, V. D. A study of the attitudes of mothers of cerebral palsied children toward child adjustment. Unpublished doctoral dissertation, American Univer., 1957.

25. MARK, J. C. The attitudes of the mothers of male schizophrenics toward child behavior. *J. abnorm. soc. Psychol.*, 1953, 48, 185-189.
26. MEAD, G. H. *Mind, self and society*. Chicago: Univer. of Chicago Press, 1934.
27. MILES, KATHARINE A. Relationships between certain factors in the home background and the quality of leadership shown by children. Unpublished doctoral dissertation, Univer. of Minnesota, 1945.
28. MURRAY, H. A. Toward a classification of interaction. In T. Parsons & E. A. Shils (Eds.), *Toward a general theory of action*. Cambridge: Harvard Univer. Press, 1951. Pp. 434-464.
29. ORLANSKY, H. Infant care and personality. *Psychol. Bull.*, 1949, 46, 1-48.
30. PARSONS, T., & SHILS, E. A. (Eds.) *Toward a general theory of action*. Cambridge: Harvard Univer. Press, 1951.
31. RADKE, MARIAN J. Relation of parental authority to children's behavior and attitudes. *Univer. Minn. Inst. Child Welf. Monogr.*, 1946, No. 22.
32. READ, KATHARINE H. Parents' expressed attitudes and children's behavior. *J. consult. Psychol.*, 1945, 9, 95-100.
33. SCHAEFER, E. S., & BELL, R. Q. Parental Attitude Research Instrument (PARI). Normative data. Unpublished manuscript. Library, National Institutes of Health, Bethesda, Md., 1955.
34. SHAPIRO, M. B. Some correlates of opinions on the upbringing of children. *Brit. J. Psychol.*, 1952, 43, 141-149.
35. SHOEN, E. J., JR. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, 39, 101-148.
36. STOGDILL, R. M. The measurement of attitudes toward parental control and the social adjustment of children. *J. appl. Psychol.*, 1936, 20, 359-367.
37. SYMONDS, P. M. *The psychology of parent-child relationships*. New York: Appleton-Century, 1939.
38. TOLMAN, E. C. A psychological model. In T. Parsons & E. A. Shils (Eds.), *Toward a general theory of action*. Cambridge: Harvard Univer. Press, 1951. Pp. 279-360.
39. WHERRY, R. J., & WINER, B. J. A method for factoring large numbers of test items. *Psychometrika*, 1953, 18, 161-179.
40. WILEY, J. H. A scale to measure parental attitudes toward certain aspects of children's behavior. Unpublished doctoral dissertation, Univer. of California, 1950.

EFFECTS OF PUBLIC AND PRIVATE SUCCESS OR FAILURE ON CHILDREN'S REPETITIVE MOTOR BEHAVIOR

ALFRED L. BALDWIN *and* HARRY LEVIN¹
Cornell University

This is a report of two studies on the conditions and consequences of the arousal of pride and shame in children. In common-sense terms, pride and shame are related to the public exposure of a person's characteristics. People are proud when their good characteristics are exposed, ashamed when their bad characteristics are revealed to the public. In its broadest sense *shame* is an emotional state that occurs when one's defects, poor abilities, or bad intentions are made public. *Pride* is an emotional state resulting from the public exposure of good abilities or good intentions. There is no implication that shame and pride are bipolar.

In these studies there was no attempt to measure directly the feelings of pride and shame. Rather, the situational factors presumed to arouse the feelings were manipulated and certain behavioral consequences were observed. Pride and shame, then, are intervening variables in the schema. The operations presumed to arouse these feelings are to have the subjects succeed or fail in public (the private condition serves as the control); and the dependent variables are the speed and accuracy of performance on simple visuomotor performances.

Although few empirical studies have explicitly investigated pride or shame, the following bodies of empirical literature are relevant to the present research: (a) the effects of failure stress on performance; (b) studies of audience effects; (c) the effects of praise and reproof on performance; and (d) studies of reward and punishment. Only the first two will be discussed here.

The effects of failure on subsequent behavior have been summarized by Lazarus *et al.* (3, 4) in their discussion of stress. The results are far from unequivocal, but they may be summarized, without too much injustice, in this way: stress results in an "impairment" of performance on both verbal and psychomotor tasks. One study, closely similar to the present

¹ This research was supported by a grant from the National Institute of Mental Health (M-901), "Pride and Shame in Children." We wish to thank Stanley Coopersmith, Mary O. Gallwey, and Elizabeth Starkweather for their help with these studies.

research, found on a symbol substitution task that failure stress increased speed and decreased accuracy (4).

Lazarus, Deese, and Osler (3) describe three conceptually distinct effects of failure: (a) a directive or cognitive effect, that is, a failure gives the *S* information about the inappropriateness of his behavior; (b) a motivational effect, which may be further analyzed into two varieties: (*b*₁) a need-persistent effect, a failure to attain a goal increases the strength of the motive to attain it, and (*b*₂) an ego-defensive effect, a failure conveys information about the person, it devalues the self, and instigates motives to restore that evaluation by succeeding, finding excuses for failure, devaluing the importance of the goal, etc.; (c) the third effect of failure—perhaps a consequence of the second—is “emotional,” that is, failure may disrupt adaptive behavior.

This analysis of the consequences of failure makes clear how the comparison of public and private failure fits into the paradigm. We hypothesize that a public failure has no different directive or cognitive effect than private failure. Furthermore, we can see no reason why a public failure should have a different need-persistent effect than private failure.

We would expect, however, that a public failure might well increase the ego-defensive effect by devaluing the self not only in the person's own eyes but in the eyes of the audience as well. Since one ego-defense is counteraction, failure might well increase the motivation to succeed and perhaps bring it to such heights that it disrupts performance.

A comparable analysis of the effects of success is not a mirror image of the analysis of failure. (a) The cognitive effects are generally opposite; *S* is informed that his behavior was appropriate; there is no reason to make any qualitative changes in the behavior. Success leads to stabilization; failure to variation. (b) The need-persistent effect is equivocal. There is no reason to suppose that the attainment of a goal necessarily reduces the motivation to attain it on another occasion unless something like satiation takes place. In general, however, success probably results in less increase in motivation than does failure. (c) The ego-defensive effect is not entirely clear either. Success enhances the evaluation of the self and probably does not instigate such intense motivation as failure. On the other hand, it may give the person a high self-evaluation to live up to and make it more important for him to succeed a second time than it was to succeed the first time.

Since we argue that, in general, the strength of motivation is less after success than after failure, the probability of its reaching disrupting intensity is lower. So general observation would indicate: people are more relaxed after success than failure.

Since certain experimental conditions in the present studies involved public performance, the research on audience effects is relevant. Dashiell (2) has reviewed this literature to 1936, and, referring particularly to his own work (1), sees a tendency for the *passive spectator group* to increase

speed and to decrease accuracy. It will be noted that a similar tentative conclusion was drawn from the results on failure stress.

If we accept these conclusions in spite of the feebleness of the supporting evidence, they support the expectation that a failure condition and an audience cumulate to produce more marked effects than failure alone. The supporting argument, however, is based upon the assumption that the effect is mediated through changes in motivation.

In order to predict an increase in speed, we must further assume either that: (a) *S*'s motivation is primarily to be speedy, secondarily to be accurate, so that an increase in motivation will then tend to increase speed at the expense of accuracy; or (b) that disruption of adaptive behavior in this task will appear in the form of faster more careless performance. The first alternative is preferable.

There may well be other effects of an audience. It may distract *S* from his work and may be an incentive to bid for attention or to ask for evaluation. In actual practice these did not seem significant features of any *S*'s behavior.

To summarize the literature related to the present studies, the results have been generally inconsistent but there is a slight balance in favor of the hypothesis that failure tends to increase speed and decrease accuracy on repetitive motor tasks and that an audience has the same effect.

EXPERIMENT A

Method

Subjects. There were 48 children, divided equally between boys and girls, membership in the fourth, fifth, and sixth grades, and high and low academic achievement as measured by the Stanford Achievement Test. The age range of the children was 8 to 12 years.

Procedure. The task given to *Ss* was to cancel all the *A*'s in a continuous story. The story was typed all in capitals, single-spaced, and was contained completely on a single page. If an *S* finished, he began again on a new page and could use as many pages as he had time for.

A single *S* was brought from the classroom to a nearby experimental room where he worked with *E*. *E* explained the task to *S*, emphasizing that he was to go as fast as he could. *S* practised cancelling *A*'s on a single line to insure that he understood the directions.

There was an "audience" and a "no audience" condition. For the former, the *E*, an adult male, sat beside *S* and watched him work. In the no audience condition, *E* left the room following the instructions, so that *S* worked alone.

Further, half of the *Ss* were given a success and half, a failure experience. The success-failure conditions were effected in the following manner: midway through the session—after 7½ minutes—*E* told *S* how much cancellation most people did halfway through the session. The

CHILD DEVELOPMENT

TABLE 1

MEANS OF NUMBER OF A'S CANCELLED DURING FIRST HALF AND SECOND HALF OF SESSION AND CHANGES

	HIGH ACHIEVERS			LOW ACHIEVERS		
	1st Half	2nd Half	Change	1st Half	2nd Half	Change
<i>Audience:</i>						
Success	171.5	156.2	-15.3	162.7	165.8	+ 3.1
Failure	157.2	170.3	+13.1	141.3	153.7	+12.4
<i>No Audience:</i>						
Success	168.8	155.2	-13.6	148.2	146.5	- 1.7
Failure	160.3	167.3	+ 7.0	150.3	157.6	+ 7.4

"norms" were based upon information from a previous experiment and were so chosen that every child in the success group was well beyond the norm and each child in the failure group far from achieving the norm. There were, consequently, the following four conditions with 12 Ss in each: (a) audience-success; (b) audience-failure; (c) no audience-success; and (d) no audience-failure. Of the 12 Ss in each condition, six were high and six were low achievers.

Results

There are two dependent variables in this experiment, the number of items cancelled and the number of omissions between the beginning and the last item cancelled. The sum of these two measures corresponds to the usual measure of speed.

In the design of this experiment, success or failure was introduced midway in the session, but the audience was present or absent throughout the session. For the study of the effect of failure and success and their interactions, the appropriate measure is change from the first half to the

TABLE 2

ANALYSIS OF VARIANCE OF CHANGES IN THE NUMBER OF A'S CANCELLED

Source	df	Mean Square	F	p
Audience	1			
Success-Failure (SF)	1	77.05		
Achievement	1	1700.17	5.94	< .05
Aud. $\times$ SF	1	337.50		
Aud. $\times$ Ach.	1	23.99		
SF $\times$ Ach.	1	10.66		
Aud. $\times$ SF $\times$ Ach.	1	360.37		
Error	40	22.05		
		286.00		

TABLE 3
MEANS OF NUMBER OF OMISSIONS DURING FIRST AND
SECOND HALF OF SESSION AND CHANGES

	HIGH ACHIEVERS			LOW ACHIEVERS		
	1st Half	2nd Half	Change	1st Half	2nd Half	Change
<i>Audience:</i>						
Success	29.5	20.7	- 8.8	39.2	27.0	- 12.2
Failure	59.2	68.3	+ 9.1	38.3	49.5	+ 11.2
<i>No Audience:</i>						
Success	34.0	29.8	- 4.2	35.5	34.0	- 1.5
Failure	19.8	36.8	+ 17.0	31.7	37.7	+ 6.1

second, obtained, in the statistical analysis, as an interaction of sessions and success-failure; but, for presentation, we will use the mean change.

The means of each half and the changes from first to second half of the session are shown in Tables 1 and 3, the analysis of variance of the changes from the first to second half of the session, in Tables 2 and 4.

On both measures the effects of success and failure are significantly different ($p < .05$, for A's; $p < .01$, for omissions). After a failure experience subjects cancel more A's ($p < .05$) and also omit more A's ($p < .05$) than they did before. A success experience tends to have just the opposite effect: it decreases the number of A's cancelled ($p < .20$) and the number of omissions ($p < .10$). If the two scores were added to obtain a speed measure, the differences would cumulate.

None of the other effects in the table reaches significance. While the success-failure difference in number of A's is significant for high achievers only, there is no significant interaction between success-failure and achievement. In the light of the other results reported here, the most likely interpretation is a general success-failure effect.

TABLE 4
ANALYSIS OF VARIANCE OF CHANGES IN THE NUMBER OF OMISSIONS

Source	df	Mean Square	F	p
Audience	1	123.76	7.76	< .01
Success-Failure (SF)	1	1846.26		
Achievement	1	33.85		
Aud. $\times$ SF	1	58.60		
Aud. $\times$ Ach.	1	17.51		
SF $\times$ Ach.	1	25.00		
Aud. $\times$ SF $\times$ Ach.	1	133.01		
Error	40	238.06		

CHILD DEVELOPMENT

TABLE 5

MEAN NUMBER OF A'S AND OF OMISSIONS BEFORE EITHER
SUCCESS OR FAILURE

	AUDIENCE		NO AUDIENCE	
	A's	Omissions	A's	Omissions
High Achievement	164.3	44.3	164.6	26.9
Low Achievement	152.0	38.7	149.2	33.6

The effect of failure resembles that found under speed instructions, i.e., speed is increased at the expense of accuracy. This result confirms findings previously reported (4). If we assume that an increase in effort would have this effect, we can interpret these findings as showing that the failure experience increased motivation and made the child try harder, while the success experience lets him relax. Remember that the child did not think of his performance as two trials; in the middle of a performance he was told he was well ahead or seriously behind the average. He had no inkling that his score would be computed for the second half alone. The success experience under these particular conditions might well give the child the feeling he could let up on the pressure a bit.

As far as audience is concerned, there are no effects. There is no evidence to support the hypothesis that a public failure is more motivating than a private one, nor that its effect is different for high and low achievers. While the differences are in the expected direction, they are not significant. There is even less evidence that public success has any different consequences from a private one.

Table 5 shows the effect of audience and achievement alone. Only the results from the first half of the session, before either success or failure were introduced, are used. Table 6 shows the analyses of variance for number of A's and number of omissions. As far as A's are concerned,

TABLE 6
ANALYSIS OF VARIANCE FOR A'S AND OMISSIONS IN FIRST
HALF OF SESSIONS

Source	df	No. A's CANCELLED			No. OF OMISSIONS		
		Mean Square	F	p	Mean Square	F	p
Audience	1	18.75					
Achievement	1	2296.34	1.25	n.s.	1530.02	2.83	< .10
Aud. × Ach.	1	27.00			3.52		
Error (within cell) ..	44	1835.77			450.19		
					541.11		

audience has no effect. There is a tendency, however, for the audience group to make more omissions.

EXPERIMENT B

The cancellation task in Experiment A utilizes skills in reading and recognizing letters in which most school children have had some relevant experiences and for which they bring some skills to the experiment. Our thinking was that a less habituated task might be more sensitive to the experimental manipulations. Besides the nature of the task, this experiment also differs from the other in the nature of the audience conditions and in the way success and failure were induced.

Method

Subjects. There were 60 girls, aged 8 to 11 years, all of whom were members of a summer day camp where the study was conducted. With the restriction that ages be approximately equally represented, Ss were randomly assigned to one of six experimental treatments.

The Task. A pegboard, 24 in. by 24 in., containing 528 pegs, was mounted on an easel at eye level to S. The pegs were of varying colors as was a circular area of the pegboard surrounding each peg. S was instructed to remove all pegs whose background was either red or blue, but to ignore the color of the peg itself.

Each S was given a practice row to insure that she understood the directions. The instructions were to remove all of the appropriate pegs as fast as she could. The session lasted for two minutes. Ss performed in two sessions on successive days.

Audience Conditions. There were the following three audience conditions: (a) No audience—for each of the two sessions E went to the other side of the room and turned her back to S immediately following the instructions. For half the Ss, E was a male, and for half, a female. (b) Adult female audience—E closely watched S while she performed during both sessions. (c) Adult male audience—same as (b), except that E was male.

Success-Failure. At the beginning of the second session, S was told that she had done very well during the previous session, that she had pulled more pegs than most of the other children (success); or that she had done very poorly, that she had removed fewer pegs than most of the other children (failure).

There were, therefore, six groups: three audience conditions, half of each of which experienced success, and half, failure.

Results

Two scores were calculated from the peg task: (a) the number of pegs removed in the two minute period, and (b) the number of pegs passed

CHILD DEVELOPMENT

TABLE 7

MEANS AND CHANGES AND ANALYSIS OF VARIANCE OF CHANGES IN
NUMBER OF PEGS FROM SESSION 1 TO SESSION 2

	S U C C E S S			F A I L U R E		
	Session 1	Session 2	Change	Session 1	Session 2	Change
No Audience	51.8	66.0	+14.2	49.5	66.9	+17.4
Male Audience	43.2	56.7	+13.5	47.3	63.6	+16.3
Female Audience ..	49.4	61.5	+12.1	47.7	62.7	+15.0

Source	df	Mean Square	F	p
Success-Failure (SF)	1	66.02	3.21	< .10
Audience	2	12.83		
SF × Aud.	2	.11		
Error	54	20.57		

over, or omitted, from the beginning to the last peg pulled. The means and changes from session 1 to 2 and the analyses of variance of the changes are presented in Tables 7 and 8.

The success-failure difference is significant at the .10 level for number of pegs removed and at the .05 level for omissions. The directions of the differences are consistent with those of Experiment A, confirming the tendency of failure to elicit a fast but careless performance.

Because we are dealing with a new skill, there are large learning effects. The number of pegs pulled increases and the number of omissions drops, for all conditions. To make the effects of the audiences and success failure

TABLE 8

MEANS AND CHANGES AND ANALYSIS OF VARIANCE OF CHANGES IN
NUMBER OF OMISSIONS FROM SESSION 1 TO SESSION 2

	S U C C E S S			F A I L U R E		
	Session 1	Session 2	Change	Session 1	Session 2	Change
No Audience	12.0	3.6	— 8.4	9.8	7.0	— 2.8
Male Audience	9.2	3.7	— 5.5	12.5	9.0	— 3.5
Female Audience ..	9.2	3.5	— 5.7	8.3	4.6	— 3.7

Source	df	Mean Square	F	p
Success-Failure	1	76.48	4.16	< .05
Audience	2	3.37		
SF × Aud.	2	10.68		
Error	54	18.40		

more apparent, we subtracted the general effects of practice, taken as the mean change over all the conditions, namely, an increase of 13.5 pegs and decrease of 4.4 omissions. The results, reported in Table 9, show more clearly the effects of success and failure.

As in the first experiment, the audience effect and the audience $\times$ success-failure interaction are in the expected direction but are not significant. Contrary to our expectation, being seen failing or succeeding does not intensify the over-all effects of success or failure. Likewise, it makes no difference to the subject whether she is watched by a male or female.

It will be recalled in Experiment A that, prior to the introduction of success or failure, the audience group was somewhat more careless in its performance than the unobserved group. An analysis of the first session only in this experiment reveals no such audience effect.

TABLE 9

ADJUSTED MEAN CHANGES IN NUMBER OF PEGS AND OMISSIONS FROM
SESSION 1 TO SESSION 2

	S U C C E S S		F A I L U R E	
	<i>Pegs</i>	<i>Omissions</i>	<i>Pegs</i>	<i>Omissions</i>
No Audience	— .7	— 4.0	+ 2.5	+ 1.6
Male Audience	— 1.4	— 1.1	+ 1.4	+ .9
Female Audience	— 2.8	— 1.3	+ .1	+ .7

DISCUSSION AND SUMMARY

The two foregoing studies report the effects either of audiences or audiences in combination with success and failure on children's performances in simple visuomotor tasks. The conclusion which appears warranted is that success and failure have appreciable effects on such behavior, but whether the performance is public or private does not.

Our expectation concerning the interaction between public-private and success-failure was that the public would intensify the reactions to success and failure, and particularly to the latter. That is, public failure would arouse ego defenses which would result in the subjects' trying to improve their performance. In neither study was there a significant interaction between success-failure and audience. Where there is a tendency for audiences to intensify the failure effect, it is on the number of omissions rather than the number of A's cancelled.

Since the search for an audience effect was the main object in this experiment, it may be appropriate to consider briefly some reasons for its absence. Assuming that there is an effect of an audience, it seems likely that both the public and private conditions were contaminated in the

CHILD DEVELOPMENT

present study. In the private conditions the experimenter was clearly going to know the results, even if he was not physically watching the behavior. The public condition was not *very* public, to say the least. In more recent experiments on other variables groups of three spectators have proved to have effects and we have devised equipment that appears to be self-scoring to create a better illusion of a truly private condition.

The effects of success and failure are strikingly consistent through both experiments. Success reduces output and leads to a more accurate performance; failure has the opposite effect. This finding appears especially trustworthy and general since it holds through several types of tasks, several experimental settings, and various groups of subjects. As we indicated earlier, there is also substantial agreement on this point in the experimental literature.

REFERENCES

1. DASHIELL, J. F. An experimental analysis of some group effects. *J. abnorm. soc. Psychol.*, 1930, 25, 190-199.
2. DASHIELL, J. F. Experimental studies of the influences of social situations on the behavior of individual human adults. In C. C. Murchison (Ed.), *Handbook of social psychology*. Worcester, Mass.: Clark Univer. Press, 1935.
3. LAZARUS, R. S., DEESE, J., & OSLER, SONIA F. The effects of psychological stress upon performance. *Psychol. Bull.*, 1952, 49, 293-317.
4. LAZARUS, R., & ERICKSEN, C. Psychological stress and its personality correlates: I. The effects of failure stress upon skilled performance. *J. exp. Psychol.*, 1952 43, 100-105.

THE CHOICE TO EXHIBIT¹

HARRY LEVIN *and* ALFRED L. BALDWIN

Cornell University

This is a report of two studies from a program of research on pride and shame in children. For our purposes, pride is defined as an emotional state that results from the public exposure of one's good characteristics, and shame as an emotional state that occurs when one's poor characteristics become public. Stemming from these definitions, our previous studies have involved subjects who succeed or fail in private or in public and the observations of certain ensuing behaviors (1).

The significance of pride and shame in guiding and controlling behavior is not limited to those circumstances in which the emotional states actually have been aroused. Rather, the anticipation of these states may itself exert potent influences, once the person has had experience with the emotions themselves. People behave in certain ways because they anticipate the approbation of some fantasied, not-present public, and they look forward to being visible to this audience. Contrariwise, the disapproval of some future audience restrains disvalued behavior.

Several distinctions are useful. Shame and pride are the feelings that result when certain acts or qualities are public. The anticipation of these states may occur at two points: (a) before any action has taken place, or (b) after some action has taken place, but before the characteristic or behavior has become public.

The present studies are concerned with the latter point. Subjects were given success and failure experiences and were then permitted to choose how visible they wanted to be to certain "supposed" audiences. Our hypothesis is that, regardless of the nature of the audience, people will want to be more visible after a success than after a failure experience.

A second parameter investigated is the nature of the audience that the subject anticipates. Many aspects of the relationships between a person and his audience may, in theory, be investigated; for example, their prestige relative to each other, their friendliness, frequency of contact, etc. In the present instances we chose to manipulate prestige. We would expect that a combination of failure and an audience of high prestige should make a person desire least visibility.

The experimental designs, then, call for the subjects to complete a task successfully or unsuccessfully, to anticipate audiences of lower, equal, and

¹ This research was supported by a grant from the National Institute of Mental Health, "Pride and Shame in Children" (M-901). We wish to thank Marianne Hales and Irene Stein for their help in collecting the data.

CHILD DEVELOPMENT

higher prestige, and to indicate how visible they are willing to be to each of the audiences.

EXPERIMENT A

Method

Subjects. The Ss were 60 children, 30 boys and 30 girls, in the fourth, fifth, and sixth grades, aged 9 to 12 years. The children were assigned at random to one of the three groups described below, with the restrictions that sex, grade level, and school achievement be equally represented in each group.

Procedure. Each S was taken individually by E from his classroom to a nearby experimental room. There he was shown a model of a small table made by gluing together strips of balsam wood. He was given a packet of pre-cut pieces of wood and a tube of glue and told to build a table as nearly like the model as he could. He was told that he could take as much time as he needed. When it was apparent that he understood the task, S was given the pre-experimental questionnaire, which is reproduced below. He read it aloud to insure that he could read and understand the words.

Now that you know about this model, what would you like to do with yours after you finish? Suppose that there will be a display of these models in your classroom.

Write *Yes* in front of all the things you would like to do with your own.
Write *No* in front of the things you would not like to do with your own.

- _____ Not show it to the class, but take it home.
- _____ Hold it up in front of the class.
- _____ Make one like it while the class watches me.
- _____ Display it in class *with* my name on it.
- _____ Not show it to the class, but just throw it away.
- _____ Display it in class *without* my name on it.

S was left to work on the model unobserved; he signalled E when he was finished. Depending on his treatment group, S was given one of three experiences at this point.

- a. *Success.* E chose one aspect of the completed model and in *objective terms* praised it. "Your model is very neat (or straight, or strong). It is much neater than most of the children in your grade have been able to do." (N = 24 in this group.)
- b. *Failure.* "Your model is *not* very neat. Most children in your grade have been able to build a neater model." (N = 24 in this group.)
- c. *Control.* E made no comment on the model. This group was included to observe the effects of simply completing the task on subsequent behavior. (N = 12 in this group.)

Each child was then given three post-experimental questionnaires, requiring him to anticipate, in turn, his own class, the kindergarten, and the high school as audiences. For each of the questionnaires the alternatives

offered S were the same as for the pre-experimental form above, with changes in wording to conform to the appropriate audiences.

Scoring the questionnaires. The task is to devise a procedure for assigning a "visibility" score for each alternative in the questionnaire. To do this we administered the visibility items to a fifth grade class under clearly pride and shame instructions. The Ss were not ones used in either of the experiments reported here. Following are the instructions given to the class of 30 children:

"Suppose you build something (like a model) that you are very proud of. You have done a very good job, better than most people in your class. What would you like to do with your model?"

TABLE I

RESPONSES OF 30 CHILDREN TO VISIBILITY ITEMS UNDER PRIDE AND SHAME INSTRUCTIONS

Item	PRIDE		SHAME	
	Yes	No	Yes	No
Show it to my parents	30	0	22	8
Put it in an exhibit for the rest of the class to see <i>with</i> my name on it	27	3	1	29
Hold it up in front of my class	24	6	3	27
Make one like it while the class watches me	18	12	3	27
Put it in an exhibit for the rest of the class to see <i>without</i> my name on it	4	26	16	14
Keep it, but not show it to anyone	0	30	14	16
Throw it away without letting anyone see it	0	30	12	18

Next, the instructions to the same group were reworded in terms of "a model you are ashamed of." The responses to the two questionnaires are summarized in Table 1.

Two items, *show it to parents* and *throw it away*, are inconsistent with the rest, i.e., preferences for these items were only slightly affected by the instructions. They probably involve other values that give them special status. The other five items could be grouped into two groups: three whose desirability was greater under pride than shame, and two for which the reverse is true.

On the basis of these results, it was decided that an individual's desire for visibility would be scored by assigning a score of +1 to a *Yes* answer to each of the following three items: Make one like it while the class watches me, Hold it up in front of the class, Display it in class *with* my name on it; and to a *No* answer to each of these two items: Display it in class *without* my name on it, Not show it to the class, but just throw it away. The possible range of visibility scores is therefore 0 to 5.

CHILD DEVELOPMENT

Reliabilities of the questionnaires. We are able to make only an indirect appraisal of the reliabilities of the questionnaires. The pre-experimental and post-experimental classroom audience questionnaires are similar. In spite of the intervening treatment, the responses to these two forms correlate .73, which is probably an underestimate of the true reliability.

Results

The visibility scores were analyzed according to two criteria of classifications: Treatment (i.e., control, success, failure) and Audience. The means and summary of the analysis of variance are presented in Table 2. Since the audience scores are correlated, the appropriate error term for it and the interaction is the interaction among Ss within treatment and audience, labeled "Error (w)" in Table 2.

TABLE 2
MEAN VISIBILITY SCORES AND SUMMARY OF ANALYSIS OF VARIANCE
ACCORDING TO TREATMENT AND AUDIENCE

	Control (N=12)	Success (N=24)	Failure (N=24)
Kindergarten			
Classroom	3.50	3.83	3.63
High School	3.42	3.42	3.12
	3.17	3.29	2.50

	df	Sums of Squares	Mean Squares	F	p
Treatment					
Audience	2	6.78	3.29	—	—
Treatment × Aud.	2	16.15	8.07	10.12	<.01
Error (b)	4	3.69	0.92	—	—
Error (w)	57	215.13	3.77		
	114	90.83	.7917		

Only the audience effect is statistically significant. The children are willing to be most visible to younger kindergarten children, next to their own classmates, and least visible to the supposed high school audience. By *t* test, each of the three audience means differs from the others at the 5 per cent confidence level or better. Though the Treatment × Audience interaction is not significant, the means in Table 2 are all consistent with the hypothesis that desire for visibility is greater after success, less under control conditions, and least following failure. Boys and girls did not differ in their responses to the experimental treatments.

In summary, the major finding from this study is that the prestige of the anticipated audience is an important determinant of how visible children want to be in front of that audience.

EXPERIMENT B

The above experiment indicated that children want to be less visible to a higher status audience than to one of classmates or younger children. Although there was a tendency for failure to create a desire for less visibility, this finding was not statistically significant.

The present study was designed to take advantage of some of the above results, principally by having each child both succeed and fail. We reasoned that the greater precision available from having each subject act as his own control might clarify the success-failure effect, and, in addition, yield information on the effects of the sequence of success-failure, or failure-success, on the degree of exposure the child will risk.

There is a further novel aspect of the second experiment. Previous studies by the writers indicated that, under speed instructions, failure leads to more rapid subsequent performances, whereas success leads to slower performances (1). The present instructions to the children building the models was "to work as carefully as possible." Our expectation was that failure would lead in the future to slower work, and that success would have the opposite effect.

Method

Subjects. The Ss were 30 children, all girls, aged 9 to 12 years, who were attending a day camp where the experiment was carried out. (It will be recalled that no sex differences were found in Experiment A, so we felt justified in working with a totally female sample.)

Procedure. The instructions, task, and success-failure manipulations were the same as in Experiment A. The changes were that each child had both a success and failure experience, in most cases, on successive days. The sequence of success and failure varied so that 15 Ss were assigned at random to one sequence, and 15 to the other.

The questionnaire was changed in several respects: (a) the pre-questionnaire was omitted; (b) the kindergarten audience was omitted; (c) the classroom audience was changed to "unit," which was the face-to-face peer group in the camp. As the higher status audience, the high school was retained.

To take full advantage of the visibility categories which evolved from the scoring check group (see above), the following categories were used. Illustrated is the pattern of responses which were scored as maximum visibility.

- | | |
|-----|--|
| Yes | Make one like it in front of the unit. |
| Yes | Hold it up in front of the unit. |
| Yes | Display it in the unit <i>with</i> my name on it. |
| No | Display it in the unit <i>without</i> my name on it. |
| No | Keep it, but not show it to anyone. |

Finally, the time the child took to complete the model was recorded.

CHILD DEVELOPMENT

Reliabilities of questionnaire responses. The reliabilities of the visibility scores may be appraised indirectly by correlating Ss' responses to the unit and high school audience questionnaires, under the same experimental treatments. The four correlations are .72, .74, .46, and .95.

TABLE 3
MEAN VISIBILITY SCORES ACCORDING TO SEQUENCE, TREATMENT,
AND AUDIENCE
($N = 15$ in each cell)

	Sequence:	SUCCESS-FAILURE		FAILURE-SUCCESS	
	Treatment:	*Success	Failure	Failure	Success
<i>Audience:</i>					
Unit		3.13	3.00	2.40	3.53
High School		2.60	2.67	2.60	3.40

TABLE 4
SUMMARY OF ANALYSIS OF VARIANCE OF VISIBILITY SCORES

	df	Sums of Squares	Mean Squares	F	p
Sequence	1	0.54	0.54	—	—
Treatment	1	7.50	7.50	8.86	< .01
Audience	1	1.20	1.20	—	—
Seq. $\times$ Treatment	1	6.53	6.53	7.72	< .01
Seq. $\times$ Aud.	1	1.63	1.63	1.93	< .20
Treatment $\times$ Aud.	1	0.53	0.53	—	—
Seq. $\times$ Tr. $\times$ Aud.	1	0.12	0.03	—	—
Error (b)	(28)	226.13	8.076	—	—
Error (w)	(84)	71.07	0.846	—	—

Results and Discussion

The two scores available for analysis are the visibility to the unit and to the high school assembly. Five is the maximum score possible to each and Audience was performed on these data. The means are presented in Table 3 and the summary of the analysis of the variance, in Table 4.

The success-failure main effect and sequence $\times$ success-failure interaction are both significant at beyond the 1 per cent confidence level. Reference to Table 3 indicates that the children are willing to be more visible with their products after success than following failure. In addition, Ss want to be most visible when success has followed a previous failure

experience. Also, when failure is the second experience, the preferred visibility is considerably higher than when failure is the first experience. In fact, in the condition where failure follows success, visibility is not much changed from the initial success experience that this group had.

There is a faint tendency for the Sequence $\times$ Audience interaction to be significant ($p < .20$). Apparently, when success comes first, Ss want to be less visible to the high school audience than to their unit, and this is true also after the subsequent failure. With initial failure the girls want to be equally invisible to both audiences and equally highly visible to both following the ensuing success.

It will be recalled that the major finding in Experiment A was that the children, regardless of success or failure, wanted to be less visible to a high school audience than to either their classmates or to kindergarten children. The appropriate comparisons for replication in the present data are the first experiences in both sequences. There is only a partial tendency to confirm the previous results. Initial success leads to a desire for greater visibility to the unit as compared to the high school assembly, but after failure, there are no differences attributable to the projected audience. One explanation is that we erred in taking the camp unit to be analogous to the classroom. The unit remains as a group for only two weeks compared to the nine-months duration of the classroom, and we actually experimented with many of the children quite early in the two-week period. If the unit members were still strange to each other, they may be as threatening an audience for poorly done work as the obviously awesome high school students. At any rate, the confirmation of the earlier findings must remain equivocal.

So far as visibility is concerned, the children are willing to be more visible after they have performed a task successfully than after failure, and they will accept greatest visibility if the success has followed a previous failure.

Success, failure and carefulness. The Ss were instructed to be as careful as possible in constructing their model and to take as much time as they needed to do a good job. They were timed from the point at which they began the construction to when they signalled *E* that they were finished. An obvious expectation is that when they previously failed they will spend more time on their next model; and that success will give them confidence in their model building talents and so permit them to hurry through.

In Table 5 are given the mean times to complete the models under the various experimental treatments. During the session following success nine of the 15 Ss spent less time building the model, whereas after failure only four children speeded up. The two sets of difference scores do not differ significantly, however ($t = 1.20$, $p > .20$). The evidence is only suggestive that failure induces more care in subsequent performances. This tendency is true only for conditions where the standard of excellence is a careful

CHILD DEVELOPMENT

performance; the evidence is ample that under speed instructions failure leads to more careless performance (e.g., 1).

TABLE 5
MEAN TIMES IN MINUTES TO COMPLETE MODELS

	S E S S I O N		
	1	2	2-1
Session 1 is Success	8.50	7.66	-0.84
Session 1 is Failure	7.67	8.77	+1.10

SUMMARY

Two studies were reported in which children chose how visible they wanted to be to audiences of various prestiges, depending on whether they had previously been praised or reproved for their performance on a model building task. In the first experiment, children wanted to be most visible to a younger audience, next to their classmates, and least visible to an older, high school audience. Whether they had succeeded or failed on the task made no difference in their preferred visibility.

In the second study, each child was given both a success and a failure experience on successive days, with the sequence varied across children. Children want to be more visible after their success than following their failure experiences. Also, the preference for the greatest exposure occurs when success has followed a previous failure.

The concepts of pride and shame are used to explain these findings. When children anticipate feelings of pride, they will choose to make themselves and their competencies visible. Conversely, the anticipation of shame disposes them to conceal themselves and their inadequacies.

REFERENCES

1. BALDWIN, A. L., & LEVIN, H. Effects of public and private success or failure on children's repetitive motor behavior. *Child Development*, 1958, 29, 363-372.

A FACTOR ANALYTIC STUDY OF CHILD-REARING BEHAVIORS

G. A. MILTON
*Stanford University*¹

Fundamental to most theories of personality development is the assumption that parents display consistent modes of behavior which serve as antecedent conditions for the behavior of their offspring. A variety of descriptive terms have been used to denote these assumed behavioral consistencies, such as "over-indulgent," "authoritarian," "permissive," "castrating," etc., but these terms do not always lead to efficient description because they may sometimes connote combinations of behavior which are not empirically consistent nor mutually exclusive.

The purpose of the present study was to derive, by means of a factor analysis, some dimensions which may underlie parental child-rearing behaviors. Factor analytic procedures have proved useful in exploratory studies of this type in other areas of psychology, and a few similar studies have been attempted in the area of child development (1, 3, 5). The discovery of consistent dimensions which underlie the child-rearing practices of parents would seem to be of the utmost importance to the understanding of the effects of parental behavior upon the personality development of the child.

PROCEDURES

The analysis presented in this report is based upon data collected by R. R. Sears and his associates (4) in the Harvard study of child-rearing practices. The subjects were 379 suburban New England mothers of children of ages five to six years. The mothers were interviewed concerning their child-rearing behaviors, and these transcribed interviews were coded for an extensive list of specific behaviors (*see* 4, Appendix B). From this list, 44 were chosen for inclusion in the present analysis. These represented a wide range of parental practices relevant to current developmental theory and at the same time provided adequate reliability of report. The scales representing these behaviors are briefly described in Table 1; more complete description of the scales, procedures and sample can be found in Sears, Maccoby, and Levin (4).

Two important considerations concerning this procedure must be kept in mind. Neither the selection of scales nor the subjects of this study are in

¹ Now at the University of Colorado.

CHILD DEVELOPMENT

TABLE I

SCALES USED IN THE FACTOR ANALYSIS OF CHILD-REARING VARIABLES

No.	Code	Description	Direction	N*	r†
1.	II-35	Severity of weaning	High: abrupt Low: gradual	347	.73
2.	II-36	Scheduling of feeding	High: rigid Low: self demand	336	
3.	II-38	Severity of mother's handling of feeding problems	High: severe pressure Low: no pressure	342	.83
4.	II-51	Severity of toilet training	High: very severe Low: not severe	349	.78
5.	II-53	Modesty permissiveness (going without clothes indoors)	High: entirely permissive Low: not permissive	330	.91
6.	II-56	Masturbation permissiveness	High: entirely permissive Low: not permissive	334	.84
7.	II-58	Sex play permissiveness (among children)	High: entirely permissive Low: not permissive	327	.81
8.	II-43	Demands for table manners	High: high demands Low: low demands	341	.70
9.	II-61	Standards of neatness, orderliness and cleanliness	High: high standards Low: low standards	355	.63
10.	II-62	Restrictions: care of house and furniture	High: many restrictions Low: few restrictions	351	.70
11.	II-65	Strictness about bedtime behavior	High: very strict Low: not strict	293	.74
12.	II-66	Strictness about noise	High: very strict Low: not strict	328	.80
13.	II-69	Restrictiveness: physical mobility	High: much restriction Low: no restrictions	328	.74
14.	III-11	Mother's response to dependency	High: strong disapproval Low: strong approval	339	.73
15.	III-18	Importance of child's doing well in school	High: very important Low: unimportant	326	.86
16.	II-73	Mother's standards of obedience (expectation)	High: instant obedience Low: no obedience expected	349	.74
17.	III-26	Demands for aggression toward other children	High: high demand Low: none	321	.80
18.	III-30	Permissiveness for aggression toward parents	High: completely permissive Low: not permissive	327	.83
19.	III-31	Severity of punishment for aggression toward parents	High: severe punishment Low: never punish	314	.76
20.	III-28	Permissiveness for inappropriate aggression toward other children	High: entirely permissive Low: not permissive	301	.70
21.	II-76	Praise for obedience	High: pays no attention Low: high praise	295	.84
22.	III-23	Praise for playing nicely	High: none Low: much praise	258	.02

* Given N refers to the average number of cases used in computing all correlations with the scale.

† Inter-rater reliabilities, corrected by the Spearman-Brown formula, are presented for all cases in which both raters made a scalar judgment.

TABLE I (continued)

SCALES USED IN THE FACTOR ANALYSIS OF CHILD-REARING VARIABLES

No.	Code	Description	Direction	N*	r†
23.	II-45	Praise for good behavior at the table	High: none Low: much praise	297	.92
24.	III-22	Permissiveness for aggression among siblings	High: very permissive Low: not permissive	273	.75
25.	III-46	Use of physical punishment	High: frequent use Low: never uses	339	.85
26.	III-36	Extent of use of rewards	High: regular use Low: never uses rewards	325	.81
27.	III-47	Use of deprivation of privileges	High: frequent use Low: never uses	342	.80
28.	III-57	Extent of use of "reasoning"	High: considerable use Low: never uses	203	.59
29.	III-60	How often threaten to punish, then not follow through	High: very often Low: never	333	.94
30.	III-66	Who disciplines, mother or father	High: always mother Low: always father	340	.95
31.	III-74	Responsibility for policies regarding children	High: father entirely Low: mother entirely	346	.78
32.	IV-13	Mother's evaluation of father	High: high esteem Low: very critical	344	.78
33.	IV-19	Dissatisfaction with the current situation	High: very dissatisfied Low: entirely satisfied	329	.79
34.	IV-22	Mother's self-esteem	High: high self-esteem Low: low self-esteem	343	.66
35.	IV-31	Mother's child-rearing anxiety	High: extremely anxious Low: none	336	.66
36.	II-23	Amount of mother's affectionate interaction with baby	High: great deal Low: none	348	.80
37.	II-24	Fun taking care of small baby	High: great deal Low: none	340	.84
38.	II-21	Mother's responsiveness to infant's crying	High: very responsive Low: unresponsive	341	.82
39.	III-73	Mother-father disagreement about child-rearing policies	High: much disagreement Low: perfect agreement	349	.78
40.	III-65	Nature of affectional bond, father to child	High: father hostile Low: father very warm	337	.83
41.	III-13	Affectional demonstrativeness, mother to child	High: very demonstrative Low: none	325	.76
42.	III-14	Extent mother plays with child just for pleasure	High: practically never Low: frequently	321	.80
43.	IV-14	How mother felt when she discovered she was pregnant	High: displeased Low: delighted	340	.94
44.	IV-15	How father felt when he discovered wife was pregnant	High: displeased Low: delighted	318	.92

* Given *N* refers to the average number of cases used in computing all correlations with the scale.

† Inter-rater reliabilities, corrected by the Spearman-Brown formula, are presented for all cases in which both raters made a scalar judgment.

CHILD DEVELOPMENT

TABLE 2

ANALYSIS OF CHILD-REARING VARIABLES: RAW CORRELATION MATRIX

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1.	15	06	11	-00	-10	-05	06	-05	04	03	07	05	04	02	-05	00	-03	10	07	-01
2.	..	-01	04	-03	-11	-07	19	08	16	07	03	12	-03	-00	-04	07	-04	02	16	17
3.	..	..	22	-07	-10	-13	16	10	18	04	05	12	08	04	09	11	-18	12	-01	08
4.	..	..	..	-30	-25	-25	20	19	18	06	17	-01	22	15	14	04	-27	14	02	11
5.	..	..	..	..	55	62	-18	-13	-32	-02	-24	-08	-20	-22	-20	15	39	-25	13	07
6.	..	..	..	..	..	64	-18	-14	-26	-01	-20	-22	-25	-25	-14	22	37	-23	11	-01
7.	..	..	..	..	..	..	-17	-11	-27	01	-26	-13	-19	-26	-12	09	39	-29	10	05
8.	..	..	..	..	..	..	..	34	39	24	17	21	03	08	22	-05	-18	14	-07	12
9.	..	..	..	..	..	..	..	..	35	25	20	08	03	12	27	-16	-29	14	-01	04
10.	..	..	..	..	..	..	..	..	..	19	34	14	12	21	31	-11	-26	22	-14	01
11.	..	..	..	..	..	..	..	..	..	..	23	06	04	03	17	-10	-08	-07	01	08
12.	..	..	..	..	..	..	..	..	..	..	..	07	15	27	24	-13	-14	11	-22	-07
13.	..	..	..	..	..	..	..	..	..	..	..	..	-02	09	03	-04	-12	07	-13	05
14.	..	..	..	..	..	..	..	..	..	..	..	..	..	16	17	-03	-16	07	-10	08
15.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	07	03	-07	14	05	-05
16.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	08
17.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	-06	-32	13	-10	05
18.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	05	15	41	-15
19.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	-47	-07	-03
20.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	12
21.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
22.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
23.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
24.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
25.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
26.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
27.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
28.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
29.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
30.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
31.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
32.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
33.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
34.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
35.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
36.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
37.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
38.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
39.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
40.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
41.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
42.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
43.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
44.	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

TABLE 2 (continued)

ANALYSIS OF CHILD-REARING VARIABLES: RAW CORRELATION MATRIX

23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44		
02	03	04	-04	-02	06	08	04	00	-11	04	-08	13	-09	01	-04	-01	-00	-08	07	-01	09		
12	-05	-04	04	06	-00	01	02	05	02	-03	-08	21	-12	-13	-26	01	-02	-00	06	-00	-04		
-10	-08	14	09	07	-13	-06	02	-05	-03	-03	-06	13	02	01	-04	16	-04	01	-08	03	-02		
-01	-14	28	09	13	-22	-01	-01	-10	-15	16	-15	19	-15	-13	-11	14	09	-08	08	08	10		
-00	12	-29	01	-04	11	-04	-09	11	15	02	14	-07	01	08	07	-11	-04	15	-09	-05	-06		
03	01	-24	04	-13	14	-13	-02	04	10	11	13	-00	10	07	09	00	07	16	-12	-04	01		
02	10	-26	10	-08	10	-06	-02	09	06	09	11	-11	02	07	01	-02	03	19	-11	-03	-05		
-03	-01	04	14	17	06	-02	-02	-03	-00	-09	06	04	-05	-04	-07	-03	02	06	-07	03	00		
-08	-00	-02	19	25	15	-15	-01	-01	16	-06	17	-07	-02	-00	-03	-10	-17	14	-15	-07	-09		
-03	-08	10	09	14	03	00	01	-11	-01	-05	05	04	02	-05	-10	01	-01	06	00	-04	-06		
16	04	-00	05	20	-00	-06	-16	00	08	-14	12	-12	-01	-10	-06	-04	-04	00	-12	-12	-13		
-08	-08	13	02	02	-06	04	-04	04	-11	-00	-10	03	-07	-09	-09	01	03	-09	05	02	01		
-02	09	00	00	08	14	-04	14	-04	01	01	06	-01	07	-05	02	-03	-10	04	-04	-13	-09		
06	-02	21	00	21	-18	12	09	-09	-15	04	-03	-04	-10	-10	-14	09	06	-14	15	-01	-04		
-00	-11	12	-01	08	-03	01	-01	-09	-03	01	-03	-05	01	06	-07	01	-08	-05	02	11	-05		
05	-30	16	19	31	-08	-09	01	-00	02	-06	06	-00	-04	-08	-12	03	-17	00	-02	-03	-10		
10	09	11	-05	14	-07	00	-10	07	-05	09	-07	06	10	-03	-02	01	12	01	05	-07	00		
-08	13	-20	-11	-29	21	-03	-01	12	03	00	-02	00	04	05	22	03	04	04	-04	10	04		
-12	-08	40	08	22	-14	06	08	-09	-08	06	-18	16	03	-00	-06	05	10	00	03	-09	00		
07	30	01	-04	00	-09	-08	-10	12	10	09	01	04	01	04	-09	-12	04	05	06	07	10		
23	-02	-01	-01	05	-01	-07	02	-05	03	03	-09	-04	00	-10	02	-09	-01	-03	08	-07	-03		
37	16	-16	-11	-20	-16	-09	-00	03	-00	-03	-05	-02	-09	-04	-01	-08	00	-18	18	09	04		
	-07	-06	-09	05	-12	04	-19	06	02	04	-04	-02	-12	-15	-12	-01	-05	-17	17	08	00		
		-00	-00	-08	01	-11	-02	08	-02	06	-16	04	-05	00	01	09	11	-04	08	15	-12		
			11	16	-33	-05	-09	-06	-10	09	-14	05	-04	-03	-01	11	12	-07	01	05	-03		
				23	00	-08	00	-00	04	00	13	-11	03	05	-10	-01	-10	14	-21	-02	-12		
					-04	-06	-01	06	-02	-07	13	-14	04	-02	-13	11	00	02	-06	-04	-10		
						-11	11	10	15	-05	15	02	12	19	10	-05	-06	19	-08	-08	02		
							06	-01	-21	14	-25	18	-01	-01	-00	21	09	-07	09	12	15		
								-34	-06	06	01	00	02	08	02	-05	-08	05	-02	04	03		
										13	-02	-02	-02	-01	00	-01	02	-09	-04	00	00	05	
											-23	16	-18	09	04	00	-56	-46	15	-07	-19	-24	
												-32	32	-11	-12	05	22	18	-10	11	17	22	
													-65	09	14	-02	-26	-16	23	-16	-16	-15	
														-04	-06	04	30	15	-07	11	07	11	
																26	22	-09	-18	29	-30	-24	-22
																	16	00	-05	15	-08	-11	00
																		-03	01	08	-12	-02	01
																			40	-06	01	17	21
																				-18	19	06	28
																					-47	-19	-06
																					09	07	
																							-52

any sense statistically "representative." The selection of these 44 scales from the total of more than 200 reflects, to an extent, the theoretical biases of the experimenters. The subject sample was chosen from two metropolitan suburbs of the north-eastern United States, and certain parent-child pairs, such as those from broken homes or those in which the child was a twin, were excluded from the sample. The dimensions which emerge from a factor analysis depend in part on the scales and the subjects who are used in the study.

Although not directly related to the factor analytic procedure, another consideration may be important. The scales used in this study refer to the mother's report of her own behavior. Little is known about the veracity of mothers' self-reports.

The extent of the interrelationship among the 44 variables was determined by computing the product-moment correlations between the score on each variable and every other variable. Scores were not available for all subjects on all scales, because some scales did not apply to all subjects and because there was insufficient evidence for making a scalar rating for some subjects on some scales. Therefore, it was not possible to compute inter-correlations using all the subjects in the sample. Table 1 indicates the average number of cases used to compute the correlation of each scale with every other scale. Since the lowest number used was greater than 200, and since there is little evidence of a systematic bias in the elimination of subjects, this condition probably does not seriously limit the results of the study.

The intercorrelation matrix was factor analyzed by the centroid method, and the resulting centroid factors were rotated by an orthogonal graphic method until what appeared to approximate simple structure was achieved. In order to reduce the effects of preconceptions, a blind rotation procedure was followed until all the structures had emerged.

RESULTS

The intercorrelations among the 44 child-rearing variables are presented in Table 2. For the most part, the degree of intercorrelation is low. This fact alone suggests that there is no strong general factor underlying the child-rearing behaviors included in this study. On the other hand, any correlation of approximately .14 or larger meets the criterion of statistical significance at beyond the .05 level, and 226 correlations meet this criterion.

Despite the large number of low order and zero order correlations in the intercorrelation matrix, inspection of the matrix indicates that there are apparent clusters of variables which seem to be mutually interrelated. For this reason, it was decided that factor analysis of the matrix might yield some stable group factors.

Seven factors emerged from the centroid analysis. After the seventh centroid factor was extracted, the estimated standard deviation of the partial

TABLE 3

ANALYSIS OF CHILD-REARING VARIABLES: ROTATED FACTOR MATRIX

Scale No.	F A C T O R S							h^2
	A	B	C	D	E	F	G	
1.	11	-14*	-03	-05	03	07	05	04
2.	09	-11	-18	01	-10	-02	37*	20
3.	22*	-06	01	16	17	-11	14	14
4.	34*	-25	-15	24	19	06	02	30
5.	-66*	11	10	09	00	-05	14	49
6.	-70*	04	19	05	-01	-11	11	55
7.	-71*	06	13	18	-08	-14	07	59
8.	37*	-02	04	32	-15	10	35	40
9.	35	19	-05	38*	-11	12	13	35
10.	48*	-01	04	32	-13	08	19	39
11.	10	14	-08	26	-20	10	33*	26
12.	38*	-13	00	21	-12	20	00	26
13.	18	08	06	-05	-05	-02	32*	15
14.	38*	-07	-36	-12	-06	04	03	30
15.	32*	-02	-01	03	05	18	-06	14
16.	32	10	-09	45*	-05	-05	02	33
17.	-23	-05	-10	06	45*	-01	19	31
18.	-51*	-09	-20	-17	-35	16	00	48
19.	43*	-13	02	15	41	-06	-12	41
20.	-30	-09	-05	-10	42*	35	21	45
21.	05	-15	04	-09	00	-05	24*	10
22.	-03	-01	-28	-38*	-01	09	08	24
23.	-07	02	-38*	-37	-06	-00	21	33
24.	-31*	-12	-03	06	-04	-21	09	17
25.	31	-15	-05	19	44*	05	-14	37
26.	04	14	09	40*	10	-04	05	20
27.	16	08	-10	51*	13	00	08	32
28.	-04	07	37*	-18	26	-11	05	26
29.	06	-35*	-05	-12	09	16	-17	21
30.	15	01	04	-08	11	-22*	-10	10
31.	-20	00	05	00	07	30*	18	17
32.	-06	56*	-01	-10	09	18	23	42
33.	-08	-47*	-01	04	15	01	-02	25
34.	-10	59*	09	34	-21	13	-08	55
35.	10	-56*	08	-17	19	16	33	53
36.	02	30	34*	-10	14	-26	09	31
37.	-06	19	26*	-05	06	-04	-09	12
38.	-15	05	27*	00	00	-07	-11	13
39.	-01	-60*	09	15	-05	-23	-06	45
40.	-02	-48*	-04	-18	-04	-02	07	27
41.	-03	28	41*	10	-04	20	21	34
42.	-04	-24	-50*	-14	02	22	-23	43
43.	-07	-42*	-15	05	-16	28	-05	31
44.	-13	-51*	06	07	-09	34	-08	42

 $M_{h^2} = 31$

* Indicates the scale's highest loading with any factor.

CHILD DEVELOPMENT

residuals had reduced to .06, thereby approximating McNemar's (2) criterion for stopping the analysis. The obtained communalities for the seven factors account for an average of 31 per cent of the total variance. This indicates that each factor accounts for an average of 4 per cent of the total variance. This small percentage is, of course, a result of the low average intercorrelation of the original correlation matrix.

The rotated factor matrix is presented in Table 3. All loadings of .30 and above have been italicized to indicate that they probably have a stable relationship to the factor. Six scales do not have a loading of .30 or above on any factor. The highest loading for a scale with any factor is indicated by an asterisk in Table 3. Those tests which do not have any loading of .30 or above will be included in parentheses in the interpretive summary of the factor with which they have the highest loading. It must be acknowledged, however, that a test with a loading of less than .30 in a particular factor has less than 9 per cent of its variance in common with that factor, and therefore is contributing little to the definition of the factor. Factor A has six loadings of .40 and above, and factor B has eight loadings of .40 and above, indicating that they are probably better defined by this analysis than the other five factors which have four or less loadings of this magnitude. Factors F and G are very poorly defined.

INTERPRETATION OF THE FACTORS

Factor A clearly seems to relate to the *strictness or nonpermissiveness of parental behavior*.

Scale	Loading
10. house and furniture restrictions	.48*
19. punishment for aggression toward parents	.43*
12. noise restrictions	.38*
14. disapproval of dependency	.38*
8. demands for table manners	.37*
9. standards of neatness and orderliness	.35
4. severity of toilet training	.34*
15. importance of doing well in school	.32*
16. standards of obedience	.32
25. use of physical punishment	.31
(3. severity of handling feeding problems	.22*)
20. permissiveness for aggression toward other children	-.30
24. permissiveness for aggression toward siblings	-.31*
18. permissiveness for aggression toward parents	-.51*
5. modesty permissiveness	-.66*
6. maturation permissiveness	-.70*
7. sex play permissiveness	-.71*

* Indicates that the scale has its highest loading with this factor.

Although the three scales which have the highest loadings with this factor suggest nonpermissiveness in the sexual area, the factor seems to deal with a much more pervasive and general content area. The same

mothers who were nonpermissive with respect to sexual behavior also tended to require the observance of property restrictions, neatness, orderliness and cleanliness, quietness, obedience, nonaggression, and expected the child to do well in school. They tended not to permit dependency and were severe in their toilet training procedures. It is interesting to note that two scales involving punishment, numbers 19 and 25, are to some extent loaded with this factor, while the scale using reasoning, number 28, does not appear related to the factor.

Factor B is less easily labeled, but it seems to be a reflection of the *general family interaction or adjustment*.

Scale	Loading
34. mother's self esteem	.59*
32. mother's evaluation of the father	.56*
36. affectionate interaction with the baby	.30
(1. severity of weaning	-.14*)
29. threaten punishment and not follow through	-.35*
43. mother's displeasure when she discovered pregnancy	-.42*
33. dissatisfaction with the current situation	-.47*
40. father's hostility toward child	-.48*
44. father's displeasure when he learned of pregnancy	-.51*
35. mother's child rearing anxiety	-.56*
39. mother-father disagreement	-.60*

This factor seems to represent a wide range of content involving the mother's attitudes toward herself, the father, and the child, the father's attitudes toward the child, and the extent of agreement between the mother and father as to the best child-rearing procedures. It appears that the mother who evaluated herself highly also had high regard for the father and that there was substantial agreement between mother and father. The mothers who scored in the positive direction on this factor also wanted the child and tended to have a consistent affectionate relationship with the child.

Factor C appears to be primarily concerned with the *warmth of the mother-child relationship*.

Scale	Loading
41. mother's affectionate demonstrativeness toward child	.41*
28. extent of use of reasoning	.37*
36. mother's affectionate interaction with child	.34*
(38. mother's responsiveness to infant's crying	.27*)
(37. fun taking care of small baby	.26*)
14. disapproval of dependency	-.36
23. praise for good behavior at table (high : none)	-.38*
42. time to play with the child (high : little time)	-.50*

If all the scales loaded with this factor are scored in the direction of a warm mother-child relationship, it becomes clear that warmth forms a consistent factor. Not only do mothers who scored high on this factor have a warmer and more demonstrative relationship to the child, but also they spend more of their time with the child just for pleasure. Interestingly, the mothers who had the warm relationship with the child also tended to

CHILD DEVELOPMENT

use praise and reasoning rather than punishment as means of control. It is not clear, however, why the table manners scale, number 23, should be loaded with this factor when the other praise scales are not.

Factor D seems to reflect a *responsible child-training orientation*.

Scale	Loading
27. deprivation of privileges	.51*
16. mother's standards of obedience	.45*
26. use of tangible rewards	.40*
9. standards of neatness and orderliness	.38*
34. mother's self esteem	.34
8. demands for table manners	.32
10. house and furniture restrictions	.32
23. praise for table manners (high : none)	-.37
22. praise for playing nicely (high : none)	-.38*

The isolation of this factor, although it is conceptually somewhat similar to factor A, seems due to the method of control which mothers who score high on this factor tend to use. Obedience, orderliness, table manners, and general good behavior seem to be valued, but these behaviors are elicited from the children by the use of rewards, deprivation of privileges, and praise. That is, these mothers seemed to employ a reinforcement theory in training their children.

Factor E presents a clear picture of the parents' attitude toward *aggressiveness and punitiveness*.

Scale	Loading
17. demands for aggression toward other children	.45*
25. use of physical punishment	.44*
20. permissiveness for aggression toward other children	.42*
19. severity of punishment for aggression toward parents	.41
18. permissiveness for aggression toward parents	-.35

The parents who had high scores on the variables with positive loading with this factor encouraged aggression, even inappropriate aggression, toward others but did not tolerate aggression toward themselves. The link between physical punishment and aggression which is found in this factor is of considerable theoretical significance. The use of physical punishment as a means of control appears related to demands for aggression so long as the aggression is not directed toward the punishing agent.

Factor F is very poorly defined.

Scale	Loading
20. permissiveness for aggression toward other children	.35
44. father's displeasure when he learned of pregnancy	.34
31. father responsible for policies	.30*
(30. mother does most of the disciplining	.22*)

This factor, if it does represent any stable set of relationships, would seem to involve a tendency for the mother to perceive the father as the controlling agent in the family.

Factor G is another rather poorly defined dimension. It may, however, seem to have a conceptual theme of the mother's concern for the child's physical well-being.

Scale	Loading
2. rigidity of the feeding schedule	.37*
8. demands for table manners	.35
11. strictness about bedtime	.33*
35. mother's child rearing anxiety	.33
13. physical mobility restrictions	.32*
(21. praise for obedience	.24*)

Whether or not these last two factors represent anything more than unstable residual factors is certainly open to question. In fact, because of the low order correlations in the original intercorrelation matrix and because of the small number of powerful factor loadings in the rotated factor matrix, the interpretation of the first five factors must be regarded as merely suggestive of significant dimensions.

DISCUSSION

Of the seven factors which emerged from this analysis, at least four are of considerable theoretical significance. Factor A, which refers to parental permissiveness, and factor C, the warmth dimension, have been identified by many investigators as important clusters of parental behavior. For example, in the analysis of Baldwin, Kalhorn, and Breese (1) and that of Schaefer and Bell (3), dimensions emerged which were quite similar to those obtained from the present study. Sewell, Mussen, and Harris (5), however, did not find a single permissiveness factor but several permissiveness factors.

The emergence of a warmth factor orthogonal to a permissiveness factor in the present study leads to the interesting conclusion that the two dimensions are independent. Therefore, it would not be surprising to find parents who were either high or low on both warmth and permissiveness in addition to the stereotyped "rejecting parent" who is cold and restrictive (1, p. 13).

The family interaction or adjustment dimension, factor B, has also been identified by other investigators but less often emerges as a separate factor. Baldwin *et al.* obtain a cluster which refers to this dimension (1, p. 7), but no corresponding factor emerges from their factor analysis. Similarly, Schaefer and Bell (3) did not isolate this dimension. What seems to have happened in both of these studies is that the family interaction or adjustment variables emerged as part of a more general "warmth" dimension. There is a similarity between factor B of the current study and the third factor of Sewell *et al.* (5, p. 144), which is identified as a "parent-child interaction factor," but factor B is more broadly defined.

Although factor E, which refers to the treatment of overt aggression, emerges as an independent dimension in the present analysis, no similar

CHILD DEVELOPMENT

factor seems to be identified in the other factor or cluster analyses reported in the literature. However, none of the other studies dealt so extensively with aggression. Factor E does have theoretical relevance in that it shows an empirical link between the use of physical punishment and the encouragement of externalized aggression in all directions except toward the punishing agent.

The relevance of factors D, F, and G is less clear than that for the other four factors. Factor D appears to be a stable dimension, and it is of some interest that the use of rewards and deprivation of privileges is linked to the mother's concern that her child be well mannered. As was pointed out previously, there may be some doubt as to the stability of factors F and G.

In general, the results of this factor analysis agree more closely with the findings of Baldwin, Kalhorn, and Breese (1) and Schaefer and Bell (3) than they do with those of Sewell, Mussen, and Harris (5), although all four studies seem to indicate that there is no *one* general dimension underlying all parental behavior. The differences between studies may be accounted for partially by the different variables studied. Just how the factors which have been isolated in this and other factor analytic studies serve as antecedents for the behavior of children is, of course, a subject for extensive research (1, 4).

SUMMARY

This study is a factor analysis of 44 parental child-rearing behaviors. Behaviors were chosen which have relevance for current theories of personality development. The data used in the analysis were obtained from interviews with 379 suburban New England mothers of children of ages five to six years. Seven factors emerged from the analysis and at least five of these were identified as stable dimensions relevant to developmental theory. The results of this analysis were compared with the results of similar studies by other investigators.

REFERENCES

1. BALDWIN, A. L., KALHORN, J., & BREESE, F. H. The appraisal of parental behavior. *Psychol. Monogr.*, 1949, 63, No. 4.
2. McNEMAR, Q. On the number of factors. *Psychometrika*, 1942, 7, 9-18.
3. SCHAEFER, E. S., & BELL, R. Q. Structure of attitudes toward child-rearing and the family. *Amer. Psychologist*, 1955, 10, 319-320. (Abstract)
4. SEARS, R. R., MACCOBY, E. E., & LEVIN, H. *Patterns of child rearing*. Evanston: Row, Peterson, 1957.
5. SEWELL, W. H., MUSSEN, P. H., & HARRIS, C. W. Relationships among child training practices. *Amer. sociol. Rev.*, 1955, 20, 137-148.

IMITATION AND THE LEARNING OF INCIDENTAL CUES BY PRESCHOOL CHILDREN

WILLIAM CODY WILSON

Harvard University

The process by which one person learns to perform an act by observing the behavior of another is usually referred to as imitation. Much of learning after infancy, especially that which occurs in a social context, is dependent upon some such process of imitation. This general phenomenon of learning by imitation has two components: one, the learning of a new response by the matching or copying of a model's response; and two, the performance of the response in an *appropriate* situation in the *absence* of the model. The first of these components has received comprehensive analysis by Miller and Dollard (2). The second component, which has received relatively little attention in the past, is the primary focus of this study.

Miller and Dollard (2) refer to the second component of learning by imitation as "independent learning" and formulate the problem as one of generalizing from the situation in which the model is present to the situation in which the model is absent. They demonstrate that such learning can take place in first grade school children when the relevant common elements are very vivid and obvious, but they do not analyze this process of generalization any further.

More recently Church (1) has defined the problem of learning by imitating in terms of "learning of incidental cues" and demonstrates that such learning does occur in rats. Rats which have learned to follow leader rats will learn, while following leader rats which are responding to a light cue, to make the "correct" response to the light cue. Church further notes that this learning by imitation is as efficient for rats as learning by trial and error.

Church's findings suggest the following hypothesis: In a learning situation in which the subject's cue for the making of a response is a

CHILD DEVELOPMENT

model's performance of that response, the subject will at the same time be learning other environmental cues which are reliable indicators of his model's behavior, but which, in the beginning, were apparently incidental to the subject's task; in the absence of his model the subject will respond to the previously incidental cue as a primary cue and will perform the response appropriately.

THE EXPERIMENT

The following experiment was designed to test this hypothesis: Two groups will learn a simple two-choice discrimination problem with color-shape cues; the control group will receive no previous training; the experimental group will receive a prior experience with a similar problem in which a model's response will be the primary cue and color-shape a secondary or incidental cue. The difference in number of trials for the two groups to learn the discrimination problem will be an index of the learning of the incidental cue occurring during the experimental group's pre-training.

Subjects

The subjects were 26 children, all members of a single nursery school in the Metropolitan Boston Area. The children were from upper middle class professional and business families with a median IQ of about 130. The ages ranged from 3 years, 7 months, to 4 years, 10 months. There were 14 boys and 12 girls; the two sexes were equally distributed between the two groups.

Apparatus

The experiment was conducted in a room approximately 9 by 25 feet with a door midway one of the long sides. At one end of the room two containers of approximately the same volume (50 cubic inches) and same height (7 inches) were placed 4 feet apart on children's chairs so that they were approximately 2 feet above the floor. One of the containers was rectangular in shape and black in color; the other was cylindrical in shape and red in color.

Procedure

Discrimination problem. A female assistant brought the subject from his regular nursery school activity to the experimental room on the pretext of playing a game. He was given the following instructions: "We are going to play a game. Here we have two boxes. I am going to hide a piece of candy in one of the boxes. You are to guess which box the candy is in. If you guess right you get to keep the candy; if you guess wrong you do not get any candy. We will play the game a number of times so you can win several pieces of candy." The subject left the room and the candy was placed in one of the containers. The subject was brought into the room

again and stood about 9 feet from the two containers. He made his choice by going to one of the containers and looking in it and picking up the candy if he made a correct choice. The subject left the room and a new trial was started.

During the series of trials the candy was always placed in the same one of the two boxes. The boxes were alternated in position in a fixed pattern from trial to trial: LRLRLRLRLR, etc. The box which was the correct cue was alternated from subject to subject. The trials were non-corrected and the subject was given a series of trials until he reached the criterion of five successive trials without error.

Experimental group pretraining. The subjects of the experimental group first learned to respond to the cue of a model's behavior. They were introduced to a similar situation with similar instructions as that described above, the difference being that the containers were identical and the female assistant played the game, too. The assistant always had her trial first in the presence of the subject and always went to the container which held candy. If the subject went to the same container, he, too, found candy. After the subject had learned to go to the same container that the model chose, the differentiated containers described above replaced the identical ones. The trials continued without interruption with the subject copying the model's response for eight trials. During this time the model was consistently going to the same box differentiated from the other by shape and color but alternated in position. At the end of the eight trials the model was called away on business and the game interrupted. When the model did not return in a short time, the game was resumed without the model as described above.

Results

The subject is assumed to have learned the problem on the first trial of the criterion series and this trial is the one reported in the results.

TABLE I

COMPARISON OF CONTROL AND EXPERIMENTAL GROUPS IN LEARNING A TWO-CHOICE DISCRIMINATION PROBLEM

	Mean No. of Trials to Criterion		Range of Trials to Criterion		Mean No. of Errors	Range of Errors
Control Group (N=13) . . .	9.3	9.3	2-20	2-20	3.8	1-9
Experimental Group (N=13)	2.2*	10.2†	1- 7*	9-15†	.7	0-3
Difference	7.1‡	.9			3.1‡	

* Number of trials responding independently of model.

* Number of trials including eight trials responding to cue of model (pretraining trials).

‡ Significant at .001 level.

CHILD DEVELOPMENT

The experimental group, as shown in Table 1, learned the discrimination problem in significantly fewer trials, after beginning to respond independently of the model, than did the control group. Furthermore, with pretraining trials in which the experimental group was responding to the model's cues included, there were no significant differences between the groups in terms of total number of trials to learning criterion. Another index of learning is number of errors made. Table 1 shows that the experimental group made significantly fewer errors than did the control group in learning the problem.

INTERPRETATION AND DISCUSSION

The difference between the two groups in learning the discrimination problem represents the amount of learning of the color-shape cues which occurred during the pretraining when the experimental group was responding to the model's behavior as the primary cue. The learning of incidental cues does occur in such circumstances and may be considered as confirmation of the original hypothesis.

It is interesting to note that when the number of trials to criterion of the control group are compared with the total number of trials to criterion, including the pretraining trials, for the experimental group no significant difference is found, suggesting that learning by imitation was as efficient as learning by trial and error. It is possible, although it is not shown by the data, that several of the subjects may have completely learned the incidental cue in *less* than the eight pretraining trials and that learning by imitation is more efficient than trial-and-error learning in terms of number of trials. The number of errors to criterion, however, might also be used as an index of learning efficiency and such data are available. The experimental group made significantly fewer errors in reaching criterion, suggesting that imitation is a more efficient means of learning—especially if errors are costly. It should be kept in mind in this discussion that there were no instructions to learn from the model—the learning was “incidental.” Instructions which served to point up the presence of incidental cues and to bring the process into conscious operation would undoubtedly increase the utility of the process which has here been hypothesized and demonstrated.

It should be noted that, although the groups are too small to draw definite conclusions, ability to solve this problem under either control or experimental conditions does not seem to be related to sex, CA, mental age, or IQ within this population.

SUMMARY

It is hypothesized that the process of learning to perform an imitated response in an appropriate situation in the absence of the model is essen-

tially that of learning of "incidental" cues. An experiment testing this hypothesis is reported. Preschool children while utilizing a model's response as a primary cue for the performing of that same response did learn, in the absence of instructions, a secondary "incidental" cue for the performance of that response. The learning of a two-choice discrimination problem for preschool children appeared to be more efficient using the imitation method than using the trial-and-error method.

REFERENCES

1. CHURCH, R. M. Factors affecting learning by imitation in the rat. Unpublished doctoral dissertation, Harvard Univer., 1956.
2. MILLER, N., & DOLLARD, J. *Social learning and imitation*. New Haven: Yale Univer. Press, 1941.

A DEVELOPMENTAL STUDY OF RACIAL AWARENESS IN YOUNG CHILDREN^{1,2}

HAROLD W. STEVENSON *and* EDWARD C. STEWART³
The University of Texas

There have been few studies with young children concerning the development of racial attitudes and awareness. Those published indicate that children as young as three and four are capable of making discriminations between the physical characteristics of Negroes and whites, and that the frequency of discrimination increases with increasing CA (1, 2, 3, 5, 6, 8, 10). The studies also indicate that racial discrimination by Negro children is dependent upon S's skin color and that among northern children both Negro and white show a preference for the physical characteristics associated with white children (5, 7, 9). These findings are of great interest. However, further information is necessary for the understanding of the development of racial attitudes. The aim of the present study is to provide information of this type. The study investigates the responses of Negro and white children from three through seven years on a variety of tests concerned with racial awareness and racial attitudes.

METHOD

Subjects

The Ss were 225 children enrolled in private nursery and elementary schools in Austin, Texas, a city with a population of 165,000 whites and 25,000 Negroes. The children ranged in age from 3 through 7 years. There were 25 white Ss of each CA and the number of Negro Ss at each CA

¹ The study was supported in part by a grant from the Hogg Foundation for Mental Hygiene of the University of Texas.

² The writers wish to thank Miss Billye McClendon, Miss Emma Lois Smith, and Mr. Robert Bell, who assisted in testing the children; and the directors and teachers of the following nursery and elementary schools in Austin, Texas, where the study was done: Tarrytown Methodist, All Saints, Jack and Jill, Jeffrey, St. Austin's, Ebenezer, Howson, Twelfth Street, Greater Mount Zion, and Holy Cross.

³ Now at Lehigh University.

CHILD DEVELOPMENT

from 3 through 7 was 23, 13, 22, 17 and 20, respectively. The groups were approximately evenly divided according to sex. All the children at the nursery schools who were within the desired age levels were included, and the elementary school Ss were chosen at random from the class roll. In general, the white Ss were of a higher socioeconomic class, based upon parent occupational level, than the Negro Ss. The schools were not integrated and the Ss resided in sections of the city where play and frequent contacts with children of the other race were impossible. The Ss' interracial contacts were assumed, therefore, to have been primarily of an incidental nature.

Materials

Four tests were constructed for this study.⁴ The purpose of the tests was to investigate children's ability to discriminate the physical differences between Negroes and whites and their racial preferences and attitudes. Great care was exercised in constructing the tests so that the materials would not be a source of bias which might distort the results. The materials were reviewed by both Negro and white individuals while they were being prepared. The tests are described below.

1. *Discrimination Test.* The test consisted of 20 10 by 20 inch cards; 12 cards were related to race and 8 were used as fillers. The filler cards depicted three objects, such as two chairs and a table; some of the drawings were in color and some were line drawings. The remaining cards were divided into three sets of four cards each. On Set 1 each card contained line drawings of the heads of three children. Two of the three children were of the same race and two were of the same sex; i.e., a Negro boy, a Negro girl, and a white girl. Set 2 consisted of the same types of line drawings as Set 1, but the figures were all of the same sex. Set 3 was identical to Set 2 except that the drawings were in color. The white children had brunette hair and the Negro children had black hair. The faces were painted either a white flesh color or a warm brown, with appropriate highlights and shadows.

Two filler cards were presented first, followed by Set 1. Half of the Ss were then tested on Set 2 followed by Set 3 and half with Set 3 followed by Set 2. The sets were separated by two filler cards.

The S was shown the first filler card and was asked to show E the picture that differed from the other two. If S made an incorrect response, he was instructed concerning the basis of correct response. The same procedure was used with the second card. The instruction, "Show me the one that is different," was repeated before each of the remaining cards was presented and a noncorrectional procedure was used.

On Set 1 S could make a correct response by choosing the child who differed either by race or by sex, while on Sets 2 and 3 a correct response

⁴ The materials for the study were constructed by Graphic Studios, Austin, Texas.

could be made only by choosing the child on each card who differed from the other by race. All of the cards were constructed so that the correct choice varied at random among the three horizontal positions (L, M, R).

2. *Doll Assembly.* Four two-dimensional masonite dolls 6 inches high were used. The dolls included a Negro boy and girl and a white boy and girl. The coloring of the dolls was similar to that of the drawings in the Discrimination Test. The girls' costumes consisted of a blouse and a skirt, and the boys', a T-shirt and shorts. The clothing was painted in light pastel shades. Each doll was divided into three pieces by divisions at the bottom of the neck and at the waist. Each section of all four dolls was interchangeable so that all degrees of consistency could be used in assembling the dolls by race and sex.

The pieces were spread in front of *S* in a random array and *S* was asked to assemble the dolls to make four children.

3. *Dolls.* The materials consisted of four small dolls made of a soft, flesh-colored plastic. Only white dolls were available and two of these were modified to create Negro dolls. One doll of each race and sex was used. The boys were clothed in identical shirts and short pants, and the girls in identical dresses.

The *S* was shown the two dolls of *S*'s sex and was asked, "Which one looks most like you?" After *S* had made a choice, the four dolls were placed in front of *S* and *S* was asked, "Which two children would you rather play with?" The linear arrangement in which the dolls were presented was varied at random among the *Ss*.

4. *Incomplete Stories.* The test consisted of seven cards in color depicting play situations. The cards showed one or two central figures, with other figures in the foreground or background. The central figure on most of the cards was drawn in a rear view and was ambiguous with regard to race. The drawings were balanced so that Negroes and whites appeared equally often on the left and right and children of the same race were not grouped together. The clothing of all figures was in pastel shades, and on each card the color of the clothing differed for each figure. The cards and the stories are described below:

Card 1. A child is shown lying on the ground. Four boys, two Negro and two white, are playing in the background. The *S* was told, "This little boy was playing in the yard. Another little boy was very mean to him and pushed him down." The *S* was asked, "Which one of these little boys do you think pushed him down?" and then, "Which little boy is going to come over and see if he is all right?"

Card 2. Two women, a white and a Negro, are seated in the foreground. Approaching them from the center of the card are two girls, one a Negro and one white. The *S* was told, "These little girls have been playing in the park. Now this little girl is looking for her mother. Which is her mother?" For half of each racial group the question concerned the Negro girl and for half it concerned the white girl.

CHILD DEVELOPMENT

Card 3. In the foreground a Negro boy and a white boy are pulling on a rope. Each wears cowboy attire and has a gun. In the background a Negro boy and a white boy are playing. The *S* was told, "These two boys are pulling on the rope, trying to see which one can get it. Which one do you think is going to win and get the rope?" After *S* made a choice he was told, "This boy (the one who wins) is going to tie up the bad man who is one of these boys. Which one do you think is the bad man?"

Card 4. A boy in the foreground is attired in hat, mittens, and jacket. In the background are a white boy and a Negro boy. The *S* was told, "This little boy is about ready to go home from school and wants a friend to go with him. Which boy do you think he will choose to go home with him?"

Card 5. A girl is seated in the left foreground. To her right is a birthday cake on a table. Behind this and to the rear are six girls, three Negro and three white, alternated from left to right according to race. The *S* was told, "This is Ann. See her birthday cake? Her mother said that she could invite three children to her birthday party. Which three girls do you think that she is going to invite?"

Card 6. A boy is seated in the foreground looking at four children, two Negro and two white, playing in the background. The *S* was told, "This little boy is afraid to go over by these children because one of the boys always hits him. Which of the boys do you think doesn't like him and always hits him?"

Card 7. In the center background a Negro boy is seated in a red wagon. A white boy and a Negro boy stand in the background. The *S* was told, "This little boy has a new wagon and wants someone to pull him. Which boy do you think he will choose to pull him?"

Procedure

The *Ss* were tested individually at school. The Negro children were tested by a Negro examiner and the white children by a white examiner. Two orders of presentation were used. The order of presentation for half of the *Ss* was the order in which the tests are described above. For the other half the Dolls and Incomplete Stories preceded the Discrimination Test and the Doll Assembly.

RESULTS

The tests proved to be of interest to the children and in general appeared to be appropriate for the age levels tested. The Discrimination Test and Doll Assembly did, however, appear to be difficult for the 3-year-old *Ss*. The pretraining on the filler cards on the Discrimination Test was protracted with these children, and although they were eventually able to make the correct choices, it is doubtful that they fully comprehended the nature of the task. In Doll Assembly it was difficult for these *Ss* to assemble the 12 pieces to form four dolls. There were few instances in which older *Ss* had difficulty in performing the tasks.

TABLE 1

MEAN NUMBER OF SEX AND RACE RESPONSES AND NUMBER OF Ss MAKING
CONSISTENT SEX AND RACE RESPONSES ON SET 1
OF THE DISCRIMINATION TEST

CA	Choices by Race (Mean)		Choices by Sex (Mean)		Consistent Race Choices (N)		Consistent Sex Choices (N)	
	W	N	W	N	W	N	W	N
3	1.3	1.3	1.3	1.9	0	0	2	0
4	1.4	1.3	1.6	1.8	0	0	2	0
5	.7	1.4	2.8	1.9	0	1	11	2
6	1.2	1.1	2.4	2.1	3	0	11	4
7	1.4	1.4	2.5	2.6	5	5	11	10

Discrimination Test

In Table 1 are presented the average numbers of choices made on the basis of race and on the basis of sex on Set 1. There were no consistent differences among the choices made by Negro and white Ss. There was no tendency for the average number of race choices to increase with increasing age, but the average number of sex choices did increase. At all age levels the average numbers of choices by sex were greater than those by race.

Statistical analysis of the data for Set 1 is facilitated by the use of number of Ss making consistent race or consistent sex responses as the measures of response (*see* Table 1). There is a significant increase in the number of Ss at each CA making consistent sex responses ($\chi^2 = 35.58$, $df = 4$, $p < .01$), and a significant difference between the proportions of

TABLE 2

MEAN NUMBER OF CORRECT RESPONSES AND NUMBER OF Ss MAKING
CONSISTENT CORRECT RESPONSES ON SETS 2 AND 3
OF THE DISCRIMINATION TEST

CA	S E T 2				S E T 3			
	Correct Responses (Mean)		Consistent Correct Responses (N)		Correct Responses (Mean)		Consistent Correct Responses (N)	
	W	N	W	N	W	N	W	N
3	1.5	2.0	1	1	1.6	1.8	1	0
4	2.9	2.4	11	4	2.8	1.8	11	1
5	3.2	2.4	16	4	3.3	2.7	16	0
6	3.9	3.1	22	7	3.9	2.8	22	3
7	4.0	4.0	25	2	4.0	4.0	25	20

CHILD DEVELOPMENT

Ss making consistent race and consistent sex responses ($t = 5.81, p < .01$). The Negro and white Ss did not differ significantly in frequency of consistent race responses, but a greater number of white, than Negro, Ss did make consistent sex responses ($\chi^2 = 5.71, p < .02$).

The results for Sets 2 and 3 are presented in Table 2. For both racial groups there is a consistent rise with increasing CA in the average numbers of correct responses. The average numbers of correct responses for white Ss are higher than those for Negro Ss. There is a significant difference between the number of Negro and white Ss making consistently correct responses on Set 2 ($\chi^2 = 12.80, df = 1, p < .01$) and on Set 3 ($\chi^2 = 31.75, df = 1, p < .001$).

Doll Assembly

The proportion of Ss in each age group assembling all four dolls correctly by race are presented in Table 3. There is a great and relatively consistent increase with age in the proportions in both racial groups. At all CA levels except 4 the proportion of white Ss assembling the dolls correctly by race was greater than the proportion of Negro Ss. An over-all evaluation of the differences between the racial groups yields a χ^2 of 5.54, which is significant at between the .01 and .02 levels of significance.

TABLE 3
PROPORTION OF Ss ASSEMBLING FOUR DOLLS
CORRECTLY BY RACE

CA	White	Negro
3	.08	.00
4	.08	.22
5	.44	.18
6	.68	.25
7	.88	.70

Dolls

In Table 4 are presented the proportions of Ss at each CA level who chose the doll of S's race as looking more like S. The proportion of own-race choices of the white Ss begins at a chance level and rises consistently through CA 6. The proportion of own-race choices of the Negro Ss also begins at a near-chance level at CA 3, drops slightly at CA's 4 and 5 and then rises at CA's 6 and 7. At all age levels the proportion of own-race choices was greater by the white Ss than by the Negro Ss. The difference between the total own-race choices in the two groups is highly significant ($\chi^2 = 25.28, df = 1, p < .001$).

In Table 5 are presented the proportions of own-race choices made when the Ss were asked to choose the dolls with whom they would rather play. The proportions begin at a chance level for both racial groups. The white Ss then show a consistent increase in proportion of own-race choices, while the Negro Ss show a consistent decrease except at CA 7. The overall proportion of own-race choices was .20 for the Negro Ss and .47 for the white Ss. The difference between these proportions is highly significant ($t = 4.50, p < .001$).

TABLE 4
PROPORTION OF Ss CHOOSING OWN-RACE DOLL
AS LOOKING MORE LIKE S

CA	White	Negro
3	.52	.43
4	.72	.33
5	.88	.33
6	.96	.53
7	.96	.85

TABLE 5
PROPORTION OF CHOICES OF OWN-RACE
DOLLS AS PLAYMATES

CA	White	Negro
3	.52	.50
4	.60	.39
5	.64	.35
6	.82	.33
7	.82	.65

Incomplete Stories

On Card 1 the Ss were asked which of four children, two Negro and two white, pushed a child down. For the white Ss the proportion of own-race choices at all ages was relatively low and did not increase with CA (see Table 6). A higher proportion of Negro Ss than white Ss made own-race choices. The proportion of own-race choices increased slightly from CA 3 through 7 for the Negro Ss, but remained relatively constant for the white Ss. The difference between the total number of own-race choices in the Negro and white groups is highly significant ($\chi^2 = 22.67, df = 1, p < .001$).

CHILD DEVELOPMENT

On Card 1 *S* was also asked which child would help the boy who had been pushed down. There were no consistent changes in response with age for either white or Negro *Ss*. The proportions of own-race choices are lower for Negro *Ss* than for white *Ss* except at CA 7; the over-all difference is not significant ($\chi^2 = 3.78$, $p > .05$).

On Card 2 there was an increase with CA in the proportion of correct identification of the child's mother for both the Negro and white *Ss*. At CA 3 the proportions are below chance, which indicates that the *Ss* may have tended to choose the mother nearer the child (the Negro mother and white child were on the left of each pair). The total numbers of correct responses do not differ significantly for the Negro and white *Ss* ($\chi^2 = .17$).

TABLE 6

PROPORTION OF OWN-RACE CHOICES OF *Ss* ON INCOMPLETE STORIES

CA	Card 1		Card 2*		Card 3		Card 4	Card 5	Card 6	Card 7
	<i>a</i> <i>b</i>				<i>a</i> <i>b</i>					
	W	N	W	N	W	N	W	N	W	N
3	.32	.65	.52	.48	.20	.22	.64	.52	.40	.35
4	.32	.56	.76	.39	.36	.72	.56	.28	.48	.35
5	.40	.45	.56	.41	.92	.73	.72	.36	.32	.64
6	.20	.71	.56	.47	1.00	.76	.72	.41	.40	.35
7	.32	.85	.60	.75	.96	1.00	.88	.45	.20	.50

* Proportion of correct choices.

On Card 3 the proportion of white *Ss* who selected the white child as the winner of the rope was high at all CA levels and showed a general increase from CA 3 through 7. The opposite trend was found for Negro *Ss*. The choices were at a chance level at CA 3 and were either at or below this level at subsequent ages. The over-all difference between the own-race choices of white and Negro *Ss* was highly significant ($\chi^2 = 19.63$, $p < .001$).

On Card 3 the *Ss* were also asked to select a badman. No consistent developmental trends were found for either racial group in the proportion of own-race choices. In general, however, the Negro *Ss* made own-race choices more frequently than did the white *Ss* ($\chi^2 = 6.53$, $p < .01$).

The *S* was asked on Card 4 to select a child who would go home with the child in the story. There is a fairly consistent increase in the proportion of own-race choices for the white *Ss*. Except for CA 3 the proportion of Negro *Ss* making own-race choices is low. The difference between the proportions of white and Negro *Ss* making own-race choices is highly significant ($\chi^2 = 24.42$, $p < .001$).

On Card 5 *S* was asked to select children to attend a birthday party. Both racial groups made approximately chance selections at CA's 3 and 4.

The proportion of own-race choices then increased at successive CA levels for the white Ss and decreased for the Negro Ss. The over-all proportion of own-race choices was .45 for the Negro Ss and .61 for the white Ss. The difference is significant at less than the .02 level ($t = 2.42$).

The S was asked on Card 6 to choose the child the central figure feared. The proportion of own-race choices for the white Ss is at a chance level, except for CA's 3 and 7. There was a general tendency for the own-race choices of the Negro Ss to increase from CA 3 to 7. The difference in the incidence of own-race choices by the two racial groups is significant at between the .01 and .02 levels ($\chi^2 = 5.50$).

On Card 7 S was asked which child the Negro child would choose to pull him. There are no consistent developmental trends in the responses of either racial group. In general, the Negro Ss made more frequent own-race choices than did the white Ss ($\chi^2 = 6.29$, $.02 < p < .01$).

DISCUSSION

The results of this study indicate that the ability to discriminate the physical differences between Negroes and whites develops rapidly during the preschool years. The 3-year-old Ss had difficulty in making discriminations of the types required, but by the seventh year the Ss were capable of making such discriminations with ease. These results corroborate those of previous studies and provide further evidence for the early development of racial awareness.

After the third year, where the responses of both racial groups tended to be made on a chance basis, the frequency of discrimination was rather consistently greater by the white children. This finding is difficult to interpret. The white and Negro Ss represented different socioeconomic groups, and, although the relevance of socioeconomic background in making such discriminations is not clear, it is possible that factors which vary with socioeconomic status, such as intelligence, experience with individual testing, and experience with the types of visual materials used, might provide the basis for the obtained differences. A serious problem is posed in comparative studies of this type because of the near impossibility of matching southern Negro and white children satisfactorily on socioeconomic status, and until more information is available, generalizations from such studies must be extended only to the groups from which the samples are drawn.

Another interpretation of this finding may, however, be made. Young Negro and white children may differ in the degree to which they respond to physical differences associated with race as being relevant in selecting a child as being "different." Negro children live in a predominantly white society. Dolls, books, magazines, and television more frequently depict white than Negro individuals. Negro children have frequent experiences with light-complexioned Negroes. In contrast, white children living in a

segregated community have less experience with Negroes and representations of Negroes. If it is hypothesized that young children perceive objects as being "different" to the degree that the characteristics of the objects deviate from those encountered in their everyday experience, it may be predicted that young white children would perceive the physical characteristics of Negroes as being relevant for selecting an individual as being different more frequently than young Negro children would perceive the characteristics of whites as being relevant. The validity of this hypothesis may be tested in further studies employing different types of test materials and subjects who differ in their frequency of contact with members of the other race.

The items related to attitudes towards self and towards race revealed a higher frequency of negative attitudes among the Negro Ss. A greater frequency of own-race rejection on the part of the Negro Ss compared to the white Ss is seen in the lower proportion of Negro Ss making own-race choices in selecting playmates, companions to go home with, and guests for a birthday party, and in selecting own-race dolls as looking more like themselves. The Negro Ss placed the Negro children in negative positions in the Incomplete Stories more frequently than the white Ss placed white children in such positions. The Negro Ss chose Negroes more frequently as being an aggressor (Cards 1 and 6), as being less likely to give aid (Card 1), as losing in a tug of war (Card 3), and as being a badman (Card 3). The surprising thing about these results is that they were obtained with Ss so young; by the ages of 4, 5, and 6 these Ss were responding in a manner which indicated not only awareness of racial differences, but also the use of stereotyped roles. A question of great interest is that of determining how children as young as these acquire such responses to race.

It is not unexpected that the study should raise many questions which cannot be answered with presently available data. It is clear from the results, however, that in order to understand how racial attitudes develop a great deal more information about the responses of young children will have to be obtained.

SUMMARY

A series of tests involving discrimination of physical differences between Negroes and whites and attitudes toward race were presented to 125 white and 100 Negro Ss between the ages of 3 and 7 years. A rapid increase in the ability to discriminate between the races was found in both racial groups between these ages. The white Ss tended to develop such discriminations at a younger age than did the Negro Ss. The Negro Ss made a lower frequency of own-race choices than did white Ss in items involving the selection of a child as a playmate, as looking most like S, etc. The Negro Ss assigned negative roles to Negro children more frequently than the white Ss assigned such roles to white children.

REFERENCES

1. AMMONS, R. B. Reactions in a projective doll-play interview of white males two to six years of age to differences in skin color and facial features. *J. genet. Psychol.*, 1950, 76, 323-341.
2. CLARK, K. B., & CLARK, M. K. The development of consciousness of self in the emergence of racial identification in Negro preschool children. *J. soc. Psychol.*, 1939, 10, 591-599.
3. CLARK, K. B., & CLARK, M. K. Segregation as a factor in the racial identification of Negro preschool children: A preliminary report. *J. exper. Educ.*, 1939, 8, 161-164.
4. CLARK, K. B., & CLARK, M. K. Skin color as a factor in racial identification of Negro preschool children. *J. soc. Psychol.*, 1940, 11, 159-169.
5. CLARK, K. B., & CLARK, M. K. Emotional factors in racial identification and preference in Negro children. *J. Negro Educ.*, 1950, 19, 341-350.
6. GOODMAN, M. E. *Race awareness in young children*. Cambridge: Addison-Wesley, 1952.
7. HELGERSON, E. The relative significance of race, sex, and facial expression in choice of playmate by the preschool child. *J. Negro Educ.*, 1943, 12, 612-622.
8. HOROWITZ, R. E. Racial aspects of self-identification in nursery school children. *J. Psychol.*, 1939, 8, 91-99.
9. LANDRETH, C., & JOHNSON, B. C. Young children's responses to a picture inset test designed to reveal reactions to persons of different skin color. *Child Developm.*, 1953, 24, 63-80.
10. STEVENSON, H. W., STEVENSON, N. G., *et al.* Social interaction in an interracial nursery school. Unpublished manuscript.



A COMPARISON OF DISCRIMINATION LEARNING IN NORMAL AND MENTALLY DEFECTIVE CHILDREN¹

BETTY J. HOUSE and DAVID ZEAMAN

University of Connecticut

Is the IQ related to learning ability in children of the same mental age? An answer to this question has been suggested by the results of previous studies. These led to the prediction that normal children would outlearn mentally defective children of the same mental age. Imbeciles with MAs of about three years learned visual discriminations poorly (5), as compared with naive monkeys (3). Since normal three-year-olds have not been found inferior to monkeys in such learning (4), the inference was made that normal three-year-olds would outlearn defectives matched with them for MA.

In another study of mentally deficient children, IQ was found to be correlated with error scores on a visual discrimination problem, after variation in MA had been eliminated by partial correlation. The range of IQs in this population of defectives was 17 to 67. Extrapolation of the correlational relationship to higher IQs also yields the prediction of superior learning in normals of matched MA.

To check the validity of these inferences, the present study offers a more direct comparison of the visual discrimination learning of normals and defectives.

METHOD

Subjects

For the normal subject population, pupils at the University of Connecticut nursery school were used. Most of these were children of the University faculty and staff. All the children in the school participated except three

¹ This research was carried out at the Mansfield State Training School, Mansfield Depot, Connecticut. The authors wish to express appreciation to Dr. Neil A. Dayton, Superintendent, and to Dr. John Cassell and Mr. Louis F. Boly, principal and assistant principal of the Longley School, for their aid and cooperation. This investigation was supported by Research Grant M-1099 from the National Institute of Mental Health, Public Health Service. Miss Alice Fielding assisted in collection of data.

CHILD DEVELOPMENT

whose MAs fell outside the range of 4-0 to 6-2 selected for the comparison. The final *N* was 30. Intelligence testing was done with the Stanford-Binet (Form L, 1937 Revision).

The sample of mentally defective children came from those attending classes at the Longley School of the Mansfield State Training School. Intelligence test scores for these subjects were provided by the Psychology Department of the Mansfield State Training School under the direction of Miss Laura Gothberg. The Stanford-Binet (1916 Revision) was used. For an earlier study (6), all available defective subjects with MAs between 2½ and 6 had been trained on the discrimination problem described below with the exception of a few who failed to cooperate. From that group, 32 who had MAs within the range of those of the nursery school population were selected for comparison.

TABLE I
CHARACTERISTICS OF SUBJECT SAMPLES

	<i>N</i>	CA (months)			IQ			MA (months)		
		<i>Mean</i>	<i>SD</i>	<i>Range</i>	<i>Mean</i>	<i>SD</i>	<i>Range</i>	<i>Mean</i>	<i>SD</i>	<i>Range</i>
Normal										
MA year IV ...	10	48.5	3.6	43-56	111.4	8.8	96-124	53.7	2.4	49-57
Defectives										
MA year IV ...	17	148.0	45.3	88-228	40.4	10.0	27-59	52.4	3.1	48-58
Normal										
MA year V	20	54.6	3.5	50-60	125.4	9.0	112-146	67.8	3.5	60-74
Defectives										
MA year V	15	136.5	33.7	99-237	51.0	7.8	36-67	65.5	4.0	60-73

Table I shows the characteristics of the normal and mentally defective groups to be compared. Both samples included boys and girls. A description of those defective subjects with MAs 2½ to 4 for whom normal controls were not available is given in a previous paper (6).

Apparatus

The apparatus was modeled after the Wisconsin General Test Apparatus (2). It consisted of a table with a sliding stimulus tray 30 in. by 12 in. with two circular food wells 2 in. in diameter centered 12 in. apart. In the center of the table, separating *S* from *E*, was a one-way screen. The sliding tray, when pulled behind the screen, was invisible to *S*. For a trial, the tray was pushed directly in front of *S*.

The stimuli consisted of painted forms cut out from ¼ in. Masonite and mounted vertically on 4 in. by 4 in. gray Masonite bases. Five forms were used, circle, square, triangle, cross and T, each having maximum height and width of two inches. Each form was repeated in five colors,

red, green, yellow, white and black, making 25 stimuli in all. Each subject was assigned two stimuli selected at random from among these 25, with the restriction that the two stimuli should differ in both color and form (for example, a red cross and white circle). Correct and incorrect stimuli were designated randomly.

Procedure

Pretraining. On the first presentation of the stimulus tray, both food cups were left uncovered with candy placed in one of them. The experimenter asked, "Can you find the candy?" and pointed if the subject failed to pick it up. On the next trial, a gray wedge was placed over one of the food cups containing candy, with the other cup left uncovered. Again, *S* was asked to find the candy and aided by the experimenter if he failed. When the subject picked up the gray wedge without prompting (usually one or two trials), discrimination trials were begun.

Discrimination trials. For a discrimination trial, *E* pushed forward the stimulus tray with two stimuli covering the two food cups so that it was directly in front of *S*. The two stimuli were the same on every trial for a given subject with one and only one always correct. The position of the correct stimulus was varied irregularly from left to right according to a Gellerman's (1) series. *E* recorded *S*'s first response as correct or incorrect. Twenty-five trials with about 15 sec. intertrial interval were given per day until a criterion of 20 out of 25 correct was reached during a single daily session. Some of the subjects were run under correction and others under noncorrection procedure. No difference was found in performance between these two procedures.

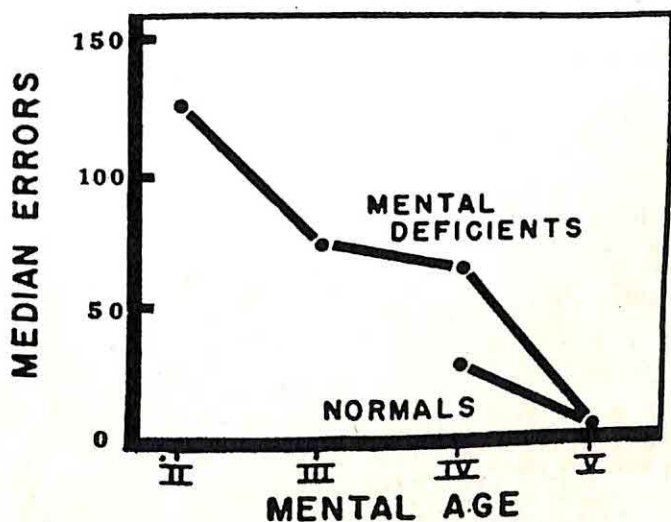


FIGURE 1. Median errors on the discrimination problem as a function of MA in years for mentally deficient and normal subjects. The roman numerals give the lower limits of the MA intervals.

CHILD DEVELOPMENT

RESULTS

As mental age increases, errors in visual discrimination decrease for both normals and defectives. Figure 1 clearly shows this. Between MA levels IV and V, the drop in errors for the defectives achieves statistical significance at the .01 level, and for the normals at the .06 level.² The non-parametric Mann-Whitney U-test (7) was used in view of the skewed distribution of error scores.

At MA level IV, the normals outlearn the defectives, a difference significant at the .05 level, while at MA V the groups do not differ. One-tailed tests of significance were used throughout because all differences were in the expected direction.

DISCUSSION

The previously inferred relationship of IQ to learning ability is supported by the superior learning ability of the normals at MA level IV. The absence of a difference at MA level V is the result of a performance limit of the task rather than equality of learning ability. Most defectives of MA V learned the discrimination in the absolute minimum of one trial. A more difficult visual discrimination would be necessary to show the expected difference at this level.

There are of course many differences between the normals and defectives other than IQ. One group is institutionalized, the other offspring of professional parents. The normals are undoubtedly healthier, happier and have richer home environments, to specify a few differences. Such differences, however, may not be sufficiently independent of IQ to allow them to be controlled and yet leave the required variation in IQ. And more importantly, it is not established that such differences are relevant variables for learning ability when IQ is controlled. The operating assumption underlying this study maintains that IQ is the major, relevant, controllable difference between the groups.

The matching of the two groups in MA may be challenged. Different revisions, 1916 and 1937, of the Stanford-Binet were used. Although there are no published comparisons of the two tests specifically for populations of mental defectives, the standardizers have shown the two tests to be as highly correlated, in a sample from the general population, as the respective reliabilities would allow (9). The standard deviations of the two tests are such however to suggest that if both groups had taken the 1916 revision the obtained difference would have increased, while if both had the 1937 revision, it would have decreased. Again, it is well to make clear another assumption, namely, that the difference in revisions is not great enough

² The over-all decrease in errors of the defective group from MA II to MA V has been found to be significant at better than the .001 level (6).

to produce a median difference of seven months in MA, that being the difference in MA necessary to produce overlap of the normal and defective functions shown in Figure 1. The matching is properly considered rough.

The findings of the present study are consistent with the proposition that defectives learn more slowly than normals of matched MA. Furthermore, these results form a fairly convincing picture when taken in conjunction with outside evidence. There are, for example, not only the experiments mentioned in the introduction whose results led by inference to a prediction of the outcome of this study, but also another experiment by Stevenson and Iscoe (8) with consonant findings. These authors report a difference in visual discrimination between normals and defectives matched for MA at the year VII level, with normals surpassing defectives in speed of learning. Finally, and in an anecdotal vein, there are the frequent reports of teachers of the mentally retarded who have long suspected that MA scores of intelligence tests were overestimating the learning ability of their students.

SUMMARY

As a check on the validity of the hypothesis, derived from previous research, that IQ is related to learning ability with MA controlled, the visual discrimination learning of 30 nursery school children of MAs IV and V was compared with that of 32 mentally defective children of approximately the same MA. The data were consistent with the hypothesis. At MA IV, the normals, with their higher IQ's, learned to discriminate significantly faster than the defectives.

REFERENCES

1. GELLERMAN, L. W. Chance orders of alternating stimuli in visual discrimination experiments. *J. genet. Psychol.*, 1933, 42, 206-208.
2. HARLOW, H. F. Responses by rhesus monkeys to stimuli having multiple-sign values. In Q. McNemar & Maud A. Merrill (Eds.), *Studies in personality*. New York: McGraw-Hill, 1942.
3. HARLOW, H. F. Studies in discrimination learning in monkeys: initial performance by experimentally naive monkeys on stimulus-object and pattern discriminations. *J. gen. Psychol.*, 1945, 33, 3-10.
4. HAYES, K. J., THOMPSON, R., & HAYES, C. Discrimination learning set in chimpanzees. *J. comp. physiol. Psychol.*, 1953, 46, 99-104.
5. HOUSE, B. J., & ZEAMAN, D. Visual discrimination learning in imbeciles. *Amer. J. ment. Def.*, (in press).
6. HOUSE, B. J., & ZEAMAN, D. Visual discrimination learning and intelligence in defectives of low mental age. (Submitted for publication)

CHILD DEVELOPMENT

7. SIEGEL, S. *Nonparametric statistics*. New York: McGraw-Hill, 1956.
8. STEVENSON, H. W., & ISCOE, I. Transposition in the feeble-minded. *J. exper. Psychol.*, 1955, 49, 11-15.
9. Terman, L. M., & MERRILL, M. A. *Measuring intelligence*. New York: Houghton Mifflin, 1937.

FACTORS RELATING TO THE ACCURACY OF ADULT LEADERS' JUDGMENTS OF SOCIAL ACCEPTANCE IN COMMUNITY YOUTH GROUPS

HELEN R. MARSHALL
University of Kentucky

Since 1950, investigators have sought reasons for the extreme variation found among school teachers in their ability to predict social acceptance within their classroom groups (1, 3, 4, 5, 6, 7). It still can only be hypothesized that factors affecting this variation in accuracy will be of three general types: (a) characteristics of individual children, (b) characteristics of groups of children, and (c) characteristics of the adult judges. This study investigated the relationship of factors of these three types to the accuracy of adult judgments of children's social acceptance. The judgments of social acceptance were made by volunteer adult leaders for members of community youth groups.

Predictions of children's social acceptance are judgments about real life events: children's demonstrated preferences of friends or companions in a particular group. Calvin (2) has attempted to shape hypotheses about the judgment process that can be tested experimentally and lead to development of theory in this area. As he points out, judgments utilize signs that indicate with varying degrees of likelihood the presence of whatever is being judged. Calvin suggests that the judge must manipulate and combine many cues to arrive at a judgment. Possibly only a few of the many cues available or used by the judge will affect the accuracy of the judgment.

The effect of only one of the many characteristics of children that might furnish cues for judgments of social acceptance has been investigated. Teachers' judgments have been reported (4, 5) to be less accurate for the six children they most and least prefer than for the class as a whole. In Calvin's wording, the teachers used their preferences as cues in the judgment process, but their preferences were not relevant to peer social acceptance of the children. The present study investigated the hypothesis that two other characteristics of the judged children, sociometric status and deviation in age from the median age of the club, relate to the accuracy of adult judgments of club members' social acceptance.

CHILD DEVELOPMENT

Characteristics of school classes have not affected the accuracy of teacher judgments of social acceptance in reported investigations (1, 3, 4, 7). This study tested the hypothesis that accuracy of leaders' judgments of youth club members' social acceptance is not affected by characteristics of the club. The three characteristics studied were size of club, median age of club members, and number of meetings held prior to the time of judgment.

Scores of adults on some measures of personality characteristics have been related to their ability to judge children's social acceptance (4, 5, 7, 11). However, Gronlund (4) has reported that differences in experience of teachers are not associated with accuracy in judging pupil's social acceptance. This investigation tested the generalization to another population of judges (4-H Club leaders) of Gronlund's negative findings for teachers with reference to two experience factors, years of experience as a 4-H Club leader and years of experience leading other community groups, and to one personality characteristic, sociometric status in an adult peer group. Additionally, this study tested the hypothesis that differences in four uninvestigated aspects of adult experience do not affect the accuracy of leaders' judgments of the social acceptance of youth club members. The four aspects of experience were membership in a 4-H Club in own youth, years of residence in the community, degree of outside club acquaintance with members, and status as a parent in regard to number, sex, and school grade level of children.

METHOD

The adult judges of this investigation were the 36 women volunteer leaders of 22 Home Economics 4-H Clubs in one Illinois county who attended the club meeting when members were given the sociometric test. The leader judges included 77 per cent of the 47 adult women leaders in the county and leaders from all but one home economics club in the county. Data for the omitted club of seven members were collected, but the leader misunderstood directions and named friends of members who did not belong to the club. The number of leaders judging social acceptance within each club varied as follows: for 11 clubs, one leader made judgments; for eight clubs, two leaders served as judges; for three clubs, three leaders served as judges.

The leaders judged the social acceptance of 263 girl members of the home economics clubs. The leaders' judgments were compared in this investigation with the club members' sociometric test score and choices of three best friends. The club members and the sociometric test procedures have been described in a report (10) concerned with members' sociometric choices. Members completed a sociometric test requesting three choices of companions from the club membership for each of three situations. Age of members was obtained from birthdates reported in 4-H Club official records.

The leaders answered a questionnaire form at the club meeting when members took the sociometric test. They were asked to name each club member's three best friends in the club. They indicated how frequently they saw each member at community gatherings other than club meetings according to these categories: four or five times a year or less, probably once a month, and more than twice a month. Their answers to check-list type questions furnished the information used in this report concerning years served as a 4-H Club leader, years of experience with other community youth groups (including Sunday school), experience as a 4-H Club member in their youth, length of residence in the community, number and school grade level of sons and daughters, and the number of meetings held by the club in the current year. Sociometric status in adult women's extension clubs (9) was obtained for 12 leaders within the two months preceding the 4-H Club tests.

Accuracy of each leader's judgment of the three best friends of each club member was determined by computing how well it agreed with the member's actual choices (unordered). Both number and percentage of agreement are used as accuracy measures in analyses of this report. A chi square test, described elsewhere (12), was used to test whether the number of agreements was beyond the .05 level of probability for the number of judgments made by each leader. The number of agreements exceeded the .05 level for 16 leaders, described as *beyond chance leaders* in this report. The number of agreements was no more than that possible at larger levels of probability for 13 leaders, described as *chance leaders*. Probability level of the number of agreements could not be determined for the seven leaders of three clubs of four, five, and six members. These leaders and clubs were omitted from all analyses using the beyond chance or chance leader dichotomy.

Accuracy of a leader's judgment of members' popularity in the club was measured by the product-moment correlation of the leader's judgment of popularity scores with sociometric scores of members in the particular club. The leader judgment of popularity score for each member was the sum of the weighted leader choices of the member as one of three best friends of other club members. Each time the leader named the member as another girl's first choice for best friend, three points were contributed to the leader judgment of popularity score for that member. The leader's selection of the member as a second best friend contributed two points. Selection as a third best friend contributed one point to the score.

RESULTS

Analyses of the number of agreements between leaders' selections and members' choices of three best friends supported the hypothesis that two child characteristics, sociometric status and deviation in age from the median age of the club, relate to the accuracy of adult judgments of youth club

CHILD DEVELOPMENT

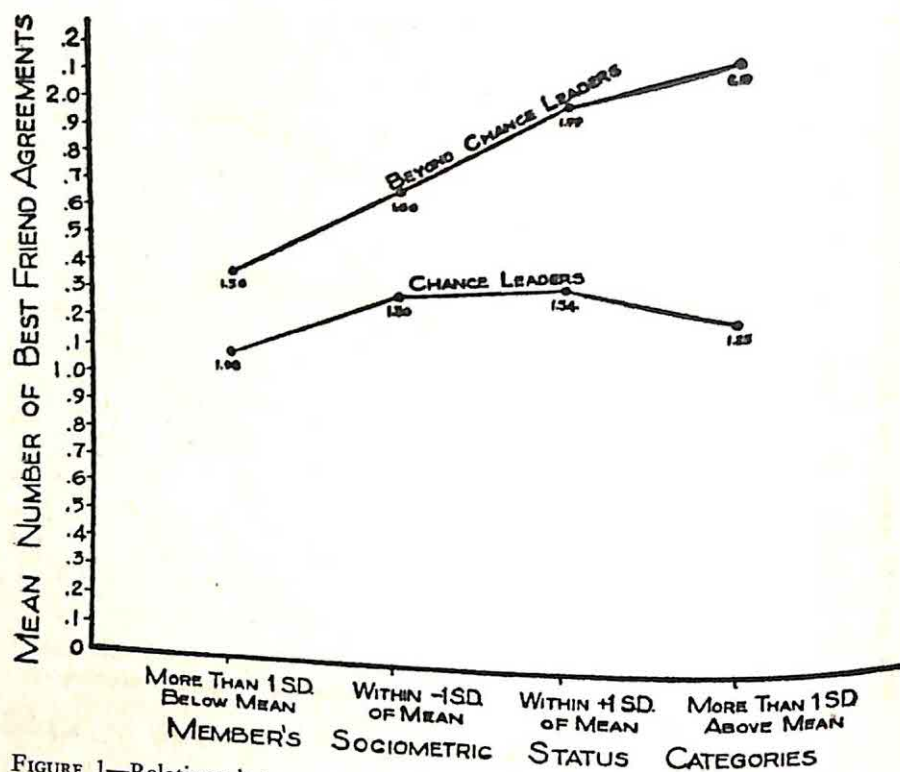


FIGURE 1—Relations between sociometric status of club members and the number of agreements obtained between club leaders' selections and the members' choice of three best friends.

members' social acceptance. The increase in mean number of agreements as the sociometric status of the member improved is shown in Figure 1 and the similar increase as the relative age position of the member improved is shown in Figure 2. Differences in mean number of agreements were significant at the .005 level for sociometric score categories and at the .05 level for age deviation categories in the two Lindquist Type 1 analyses of variance (8) used to test these and other differences. All sociometric score categories differed significantly by *t* tests, while neighboring category differences were not significant. As described elsewhere (10), the mean sociometric score of all club members was 19 points and the standard deviation, used to establish categories, was 11.2. Approximately equal numbers of members were found in the three categories for age deviation from the median age of the club listed in Figure 2.

The greater influence of members' position on the accuracy of judgments made by beyond chance leaders than those made by chance leaders, indicated in Figures 1 and 2, gave additional support to the child characteristics hypothesis. Differences in number of agreements for the two groups of

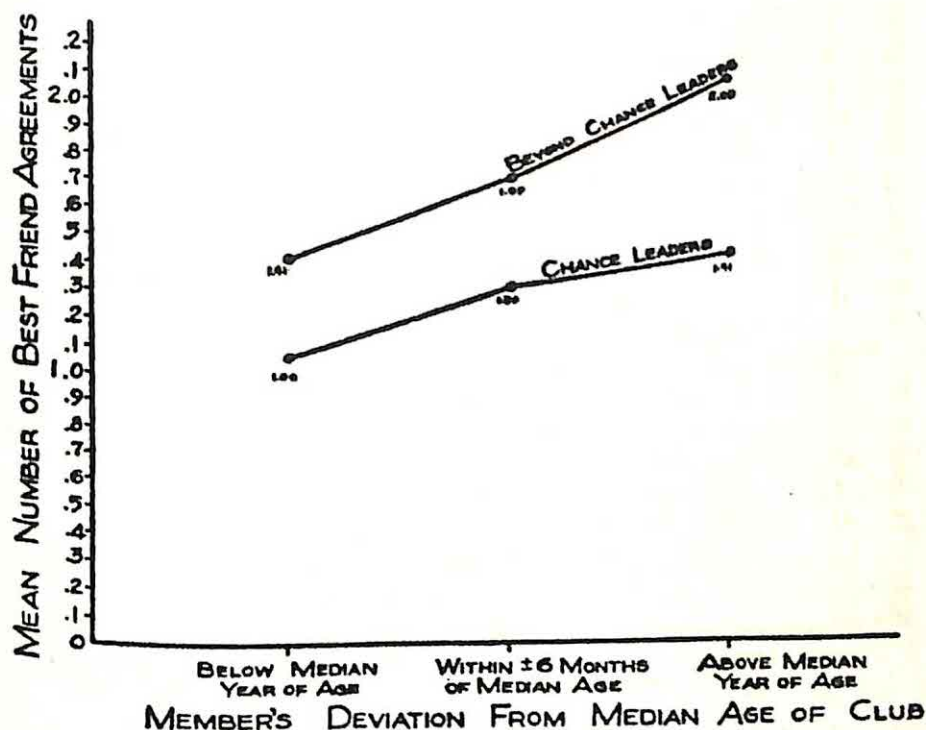


FIGURE 2—Relations between club member's deviation in age from the median age of the club and the number of agreements obtained between club leaders' selections and the members' choice of three best friends.

leaders were significant at the .005 and .001 levels in the sociometric status and age deviation Type 1 analyses of variance respectively. In simpler treatments by subjects analysis of variance tests for each group of leaders, differences between sociometric score and age deviation categories were significant for beyond chance leaders but were not significant for chance leaders. By additional tests of interaction and mean differences, beyond chance leaders improved accuracy of judgments more as the members' relative age position improved than did the chance leaders. The test of interaction between leader group and sociometric status category was not significant.

The hypothesis that accuracy of leaders' judgments of youth club members' social acceptance is not affected by characteristics of the club was rejected by the data of this investigation. Accuracy of the club leaders' judgments decreased as the size of the club increased from four to 27 members. The product-moment correlation between number of club members and percentage agreement accuracy of leader judgments was $-.64$, a large and significant coefficient. Percentage agreement of leader's selections with members' choices of three best friends ranged from 25 to 85 per cent.

CHILD DEVELOPMENT

Size of club also correlated negatively and significantly ($-.53$) with the accuracy measure for adult judgment of popularity status within the club. The range of correlations between leader judgment of member's popularity scores and sociometric scores was $-.16$ to $+.96$.

Findings for the other two club characteristics, however, support the club characteristics hypothesis. Median age of the club did not correlate significantly with the accuracy measures for leader judgments of best friends ($r = .26$) and of popularity status within the club ($r = .21$). The range of median age for the 22 clubs was 11-6 to 15-2 years. By t test, the mean number of club meetings did not differ for the 16 beyond chance leaders ($M = 7.25$) and the 13 chance leaders ($M = 7.30$).

The data of this investigation supported the hypothesis that differences in experience of club leaders do not affect the accuracy of their judgments of social acceptance of youth club members. The null hypothesis could not be rejected in any t tests of differences or chi square tests of independence for beyond chance and chance leader judges in the following six adult experience factors: years of experience as a 4-H Club leader, years of experience as a leader of other community youth groups, membership in a 4-H Club in own youth, years of residence in the community, degree of outside club acquaintance with members, and status as a parent in regard to number, sex, and school grade level of children. Neither was the adult personality characteristic of sociometric status in an adult peer group associated with accuracy of adult judgments of children's social acceptance.

DISCUSSION

Two child characteristics, sociometric status and deviation in age from the median age of the club, were shown to be associated with differences in accuracy of adult judgments, or to provide cues related to this accuracy. These findings indicate that adults differ in the cues they use to judge children's social acceptance, a difference that Calvin (2) has hypothesized will be demonstrable for the process of learning to judge.

Accuracy of these leaders' judgments depended on one group characteristic, size of the club. Studies of school classes (3, 4, 7) have not revealed significant relations between size of class and accuracy of teacher judgments of social acceptance, although all report the direction of relationship to be negative, as was found in this study. School classes studied have ranged in size from 17 to 43 pupils. It is possible that this relationship is more marked for the more disparate and smaller group sizes of this investigation. It is possible, also, that accuracy of leaders' judgments, necessarily based on fewer hours of observation than is possible for teachers, is affected more by the size of the group than is the accuracy of teachers' judgments. The nonsignificant but positive r 's found between median age of the group and judgment accuracy are in line with findings of studies of teacher judgments (1, 7).

Experience factors did not affect the accuracy of judgments of these community group leaders. This finding corroborates with a different population the findings of investigations using school teacher judges. The conclusion appears warranted, then, that experience factors of the sociological type used in these studies do not indicate how accurately an adult will judge children's social acceptance.

SUMMARY

This study investigated factors affecting the variation in accuracy of adult judgments of children's social acceptance. Adult judges were 36 women serving as volunteer leaders of 22 Home Economics 4-H Clubs with 263 girl members. Accuracy of judgments was measured by number or percentage of agreement of the leader's selections with club members' choices of three best friends and by the product-moment correlation for each club member of the leader judgment of member's popularity scores with the members' sociometric scores.

Two characteristics of the judged children affected the accuracy of these club leaders' judgments of social acceptance. Accuracy of judgments increased as the members' popularity status improved and as the members' relative age position shifted from below to above the median age of the club. The more accurate leader judges showed a greater increase in accuracy of judgment as the members' sociometric status and relative age position improved than did the less accurate judges.

Accuracy of the leaders' judgments decreased as the size of the club increased from four to 27 members. No association with accuracy of judgments was found for two other group characteristics: median age of the club and number of club meetings held prior to the time of judgment.

Differences in experience of these club leaders did not affect the accuracy of their judgments of the social acceptance of club members. Additionally, sociometric status of the leader in an adult peer group was not associated with accuracy of judgment.

REFERENCES

1. AUSUBEL, D. P., SCHIFF, H. M., & GASSER, E. B. A preliminary study of developmental trends in socioempathy: accuracy of perception of own and others' sociometric status. *Child Developm.*, 1952, 23, 111-128.
2. CALVIN, J. S. Learning to judge—an exercise in probability. Address to Univer. of Kentucky chapter of Sigma Xi, 1957.
3. GAGE, N. L., LEAVITT, G. S., & STONE, G. C. Social perception in the classroom. Urbana: Bureau of Educational Research (Mimeo.), 1955.

CHILD DEVELOPMENT

4. GRONLUND, N. E. The accuracy of teacher's judgments concerning the sociometric status of sixth grade pupils. *Sociometry*, 1950, 13, 197-225; 329-362.
5. GRONLUND, N. E. The relative ability of home-room teachers and special subject teachers to judge the social acceptability of pre-adolescent pupils. *J. educ. Res.*, 1955, 48, 381-391.
6. GRONLUND, N. E. Generality of teachers' sociometric perceptions: relative judgment accuracy on several sociometric criteria. *J. educ. Psychol.*, 1956, 47, 27-31.
7. GRONLUND, N. E. The general ability to judge sociometric status: elementary school teachers' sociometric perceptions of classmates and pupils. *J. educ. Psychol.*, 1956, 47, 147-157.
8. LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. New York: Houghton Mifflin, 1953.
9. MARSHALL, HELEN R. Some factors associated with social acceptance in women's groups. *J. Home Econ.*, 1957, 49, 173-176.
10. MARSHALL, HELEN R. Prediction of social acceptance in community youth groups. *Child Developm.*, 1958, 29, 173-184.
11. MARSHALL, HELEN R. Relations between personality characteristics of college students and the accuracy of their judgments of children's social acceptance. *J. Home Econ.*, 1958, 50, 207-212.
12. McCANDLESS, B. R., & MARSHALL, HELEN R. A picture sociometric technique for preschool children and its relation to teacher judgments of friendship. *Child Developm.*, 1957, 28, 139-147.

A CHILD STUDY LABORATORY ON WHEELS^{1,2}

SIDNEY W. BIJOU
University of Washington

A laboratory on wheels has been demonstrated to be a practical way of conducting experimental studies with children. Since it can be hauled anywhere, normal and deviant children of all ages are accessible for observation. Compared with the other ways of obtaining subjects (for example, having parents bring their children to a clinic, establishing and maintaining laboratory schools, or going to a school or institution and seeing children under make-shift conditions), observing children at the "sources of supply" in a trailer is a relatively simple, inexpensive, and flexible method. Furthermore, a unit of this type has the advantages of a laboratory in a building, i.e., observations may be made under relatively constant conditions, and apparatus for controlling conditions and recording responses may be used with convenience.

During the past four years most of the experimental studies on children at the University of Washington has been done in a laboratory-trailer. This paper describes its uses.

Figure 1 is a photograph of the laboratory parked beside a private day care center in Seattle. The masonite structure is 19 feet long by 8 feet wide by 7 feet high and it rides 14 inches above the ground. Figure 2 is the floor plan of the interior. It is the same as the laboratory in the Institute of Child Development (1). The layout consists of a large "playroom" and a small observation room with a one-way mirror. Permanent equipment in addition to the viewing mirror consists of a microphone and speaker, a heater-fan, four lights, and two juvenile tables (one built over each wheel), and chairs. The other things in the diagram are for the study in progress dealing with

¹ The development of the methods and procedures for the laboratory-trailer was supported by a grant from the National Science Foundation (G-2141).

² I wish to express my gratitude to Robert Sharply for his assistance in this work.

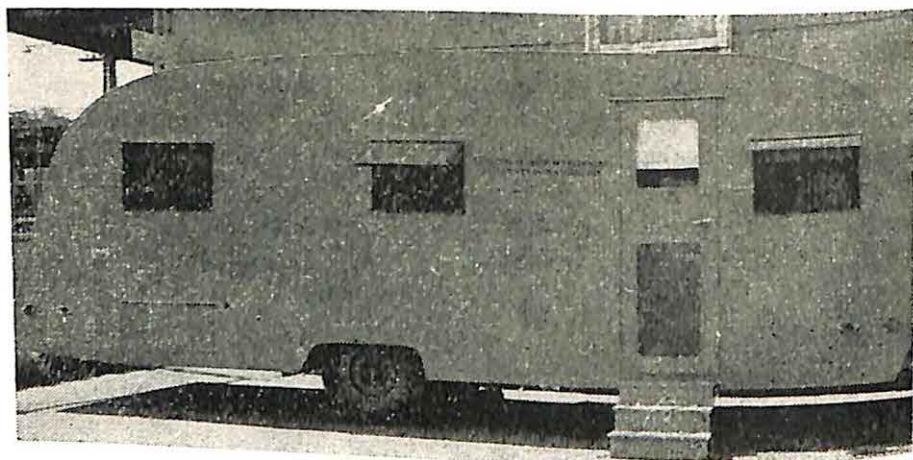


FIGURE 1—The mobile child study laboratory in operation at a private day care center.

the development of operant discrimination. They include an adult-sized chair with a right-sided writing arm, an opaque screen, a standard toy on one table, and two pieces of apparatus on the other. The one on the left contains a lever and two jewel lights; the one on the right is a device for presenting stimuli on cards. The room has a playroom atmosphere created by the beige walls, multicolored matching wallpaper on the screen, frosty lights, yellow window valences, natural colored grass rug, and augmented by the children's furniture and the bright colors of the toy and apparatus. On the table in the observation room are a control panel and two cumulative recorders. A universal dispenser and chute are attached to the partition.

The following is a typical sequence of events in the operation of the trailer:

1. After plans for a study have been completed, the trailer is brought from the last place of operation to the campus for installation of equipment and repainting. Movement is by truck from the Buildings and Grounds Department of the University.
2. Arrangements are made to locate the trailer where the study is to be performed.
3. The trailer is moved to its new location and situated close enough to a building to attach an extension cord. While one person is preparing the laboratory for operation (suspending it on jacks, checking out equipment, and the like), the other member of the team, usually a young woman, is becoming acquainted with the personnel and children. This part of the procedure takes one or two days depending upon the group.
4. To keep all the youngsters happy, the children in the class or school (if small, like a typical nursery school) who have not participated in the study are given a "turn" to play and to receive some of the reinforcers.

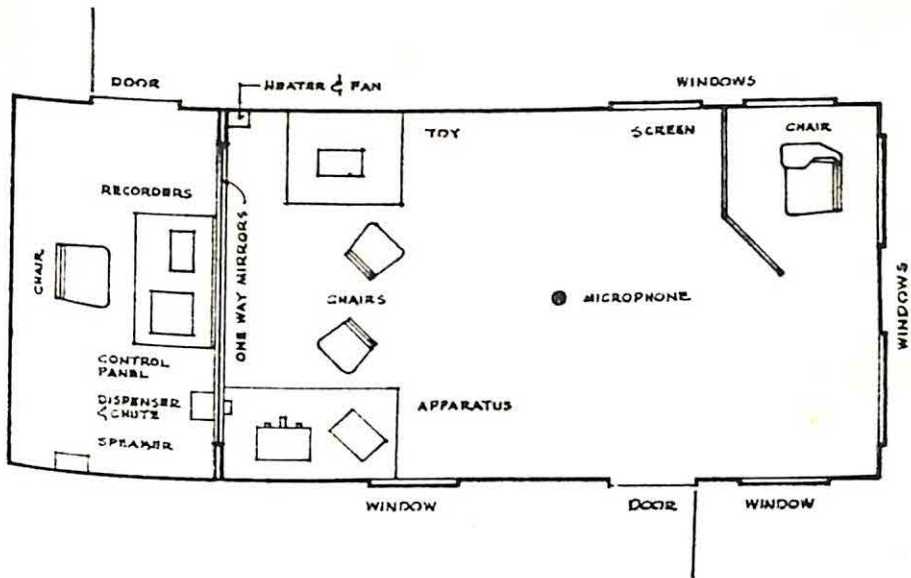


FIGURE 2—Floor plan of the mobile child study laboratory.

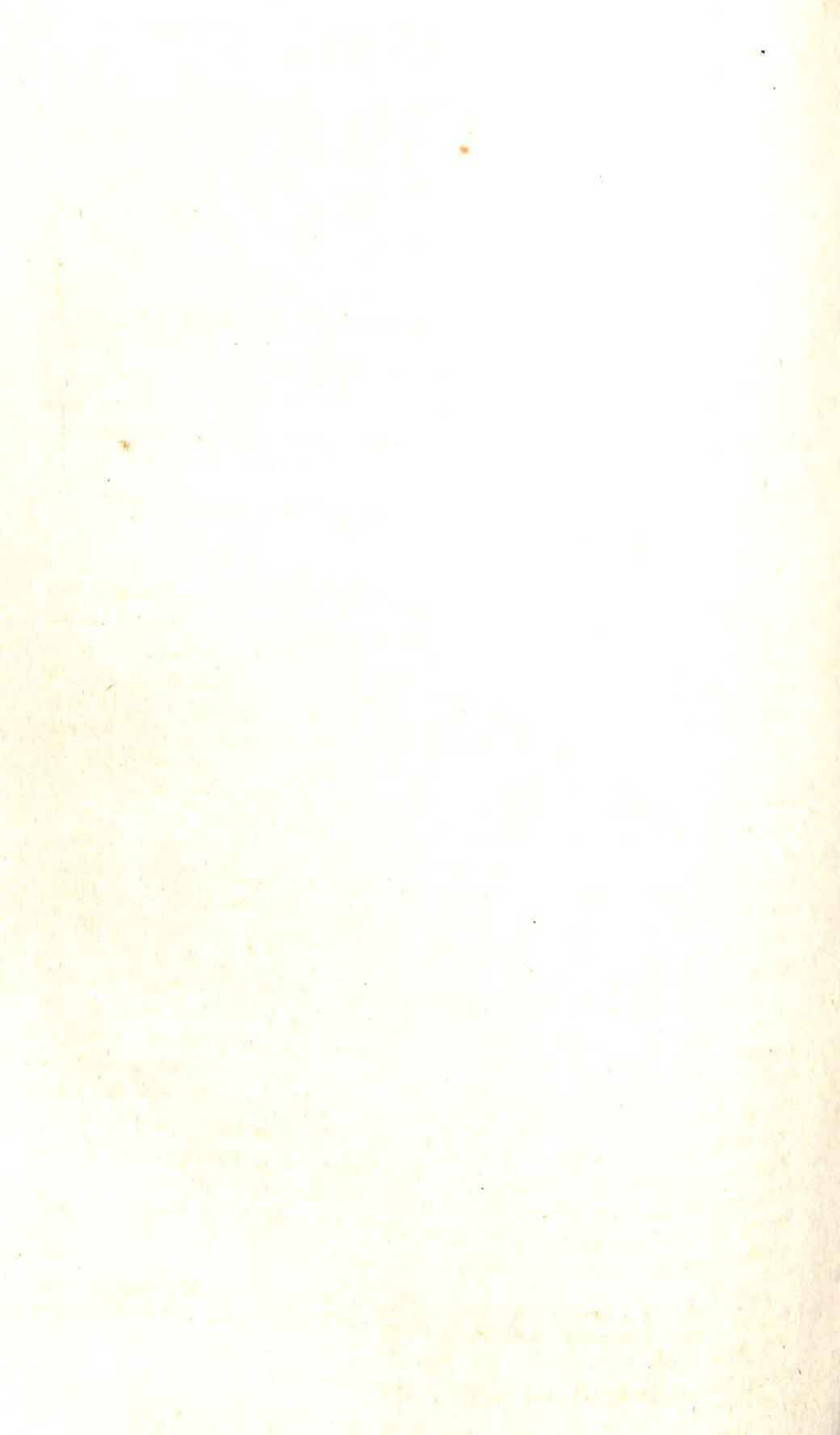
To date the laboratory has been used as an adjunct to a private day care center, a public school for exceptional children, the University student housing village, and a remedial reading clinic. Reactions of the children (especially the younger ones) have been enthusiastic. Comments from teachers and administrators have been to the effect that research conducted in the trailer fits in well with school routines and that the activity is enjoyed by the children.

Variations on the trailer described readily come to mind. For example, a longer trailer, one which would house two "playrooms" with a common observation booth between, might be more efficient in some situations. Or more elaborate permanent equipment (tape recorders, motion picture cameras, photographic cells, etc.) might be installed to obtain continuous records of activity level, activity changes, and verbalizations. Or the interior might be altered to meet specifications of an age group such as infants or adolescents, or for analysis of social behavior under free or restricted conditions. A vehicle of this sort can be tailored to almost any research problem.

The trailer was obtained by the Surplus Properties Department of the University and was renovated by the Buildings and Grounds Department for about \$900.00. Maintenance expenses consist of fees for license plates, charges for hauling, and periodic repainting.

REFERENCES

1. BIJOU, S. W. Methodology for an experimental analysis of child behavior. *Psychol. Rep.*, 1957, 3, 243-250.



THE DEVELOPMENT OF SOCIAL COMPLIANCE IN YOUNG CHILDREN¹

VAUGHN J. CRANDALL, SONYA ORLEANS, ANNE PRESTON,
and ALICE RABSON

Fels Research Institute for the Study of Human Development

Social conformity as an aspect of mental health has increasingly concerned social scientists, professional workers in the field of mental health, and laymen alike. In the last few years, observers of the American scene have noted a number of trends toward such conformity; and many of these observers have predicted dire consequences of these trends, both for individual mental health and for our general cultural progress. The determinants and processes entering into social conformity or nonconformity are obviously involved and complex. Only recently have some of the more simple relationships been understood. From a sociological point of view, a number of writers and researchers are gradually establishing a more concise picture of certain trends toward social conformity in this country, and have elucidated several factors producing some of these changes. However, much less is known about the effects of these trends on individuals who experience them.

Certain aspects of social conformity are obviously necessary for the maintenance of society. Mature members of any society are required to comply with many social demands, prohibitions, and regulations. Effective social participation, however, also demands autonomous action as well as social compliance. Socially-contributive individuals must be able to make decisions on their own, to engage in self-initiated activities, and to withstand social pressures when these run counter to personal values. Thus, extremes of conformity or nonconformity may not only limit an individual's social effectiveness but also be deleterious to his personal integration and mental health.

¹ This article is a condensed version of a paper presented by the senior author at the Fourth Annual Conference on Social Psychology and Mental Health, Vanderbilt University, April, 1958.

CHILD DEVELOPMENT

One aspect of social conformity has received little attention from psychological researchers to date, namely, the question of how social conformity or nonconformity develops in children, and the role of parents in this development. The present study was concerned with this general question.

METHOD

The sample of the study was comprised of 59 children from three through eight years of age. The children were members of the Fels Research Institute's longitudinal research population, and resided in homes within a 30-mile radius of the Fels Institute in southwestern Ohio. Almost all of the families of the children were middle class, and the parents were better educated and tested higher on intelligence tests than national averages.

The children were studied in two separate groups. A younger group was comprised of 30 children, three, four and five years of age. This group attended the Fels Experimental Nursery School during the regular school year, and will be referred to as the Nursery School children for the remainder of this discussion. A second group of children, composed of 29 six-, seven-, and eight-year-olds, attended the Fels Day Camp in early summer after their regular schools had closed. This group of older children will be designated the Day Camp children.

During the Nursery School and Day Camp sessions, the children's interactions with peers and with teachers were observed. Two full-time observers made daily time-sample observations of each child. From these observations, each observer independently rated the children in respect to a number of social behavior variables. Rating scales were specially constructed for these ratings. Their format was patterned after the Fels Parent Behavior Rating Scales, described in detail elsewhere (1, 2, 3). Each rating scale consisted of a verbal definition of the variable to be rated, and descriptive cue-points along a 90-millimeter line so that graphic subject-to-subject comparisons could be made. The central variable investigated in the study was face-to-face compliance. The rating scale for this variable was entitled "Compliance with Commands and Suggestions for Others," and was concerned with the frequency and alacrity with which the children acceded to the demands and suggestions of other persons. Other social behavior variables which were rated were: (a) "Amount of Help-seeking," (b) "Amount of Emotional Support-seeking," (c) "Amount of Approval-seeking," (d) "Amount of Hostile-aggressive Behavior," (e) "Amount of Attempts to Direct and Influence Others," (f) "Amount of Independent Achievement Efforts," and (g) "Readiness to Withdraw from Threatening Situations." With the exception of the last two variables, separate ratings were made of behaviors exhibited in peer interactions and of behaviors employed in interactions with adults.

Subsequent to the observations and ratings, the two observers, using pooled judgments, also made a 75-item Q-sort of each child's characteristic

social behavior. For each child, the two observers jointly sorted Q-sort items into seven categories running from those behaviors which were highly characteristic of the child to those behaviors which were not. A forced-normal distribution was employed in this sorting.²

Each child's social behavior was thus assessed normatively and ipsatively. The rating scales provided *normative* comparisons of each child's social behavior with that of the other children in his group; the Q-sort of each child provided *ipsative* information regarding the characteristic social behaviors of that child *per se*. This latter evaluation was not contingent on the child's relative standing in comparison to other children.

Previous to the children's participation in the Fels Nursery School or Day Camp, these children were also observed as they interacted with their mothers in their homes. These observations were made as part of the routine Fels Home Visit Program. The Fels Home Visitor, an experienced clinical psychologist, observed mother-child interactions, and subsequently rated the children's compliance toward their mothers using the same scale employed by the Nursery School and Day Camp observers to rate the children's behavior toward adults at Fels. Finally, the Home Visitor also rated the mothers' behavior for: (a) the degree that each mother rewarded her child's compliance, and (b) the amount of punishment which the mother employed for noncompliance. These latter ratings were made with rating scales whose format was similar to the rating scales previously discussed.

The data obtained from the various ratings and Q-sorts provided information relevant to the following questions: (a) What personality attributes are associated with social compliance and noncompliance in young children? (b) Do these personality correlates change with age? (c) How consistent is young children's social compliance across people and social situations? (d) To what degree are maternal reactions to children's social compliance predictive of the children's social compliance outside the home? (e) Finally, what implications do the over-all findings of the study have with respect to social conformity and mental health in young children?

RESULTS AND DISCUSSION

Degree of social compliance was not a function of either sex or intelligence in the children studied. In neither the Nursery School children nor the Day Camp children were the girls more compliant than the boys, nor were the compliant children at either age more or less intelligent than the less compliant children.³

² Mimeographed copies of the rating scales and Q-sort items may be obtained from the authors.

³ Nonparametric statistics were employed for all data analyses. Unless otherwise specified, the rank-difference correlation was used as the measure of association, and the Mann-Whitney U Test was employed to test differences between groups.

CHILD DEVELOPMENT

In the Nursery School children peer-compliance was correlated .51 with adult-compliance, significant at the .01 level of confidence. The magnitude of this correlation suggests that a number of these children who were compliant or noncompliant with peers behaved similarly with adults. However, since the relationship was far from a perfect one, it is also evident that others of these children who were compliant with peers were not so with adults, and *vice versa*. In the light of this finding, the age of the persons with whom the children did or did not comply cannot be ignored. Table 1 presents correlations of the children's other social behaviors with their peer-compliance and with their adult-compliance. In Table 1, the column to the

TABLE 1
RELATIONSHIPS OF NURSERY SCHOOL CHILDREN'S COMPLIANCE
TO OTHER SOCIAL BEHAVIORS

Variable	INTERACTION WITH PEERS		INTERACTION WITH ADULTS	
	Rho	p	Rho	p
1. Instrumental help-seeking	.37	.05	.08	ns
2. Emotional support-seeking	*	.01	.10	ns
3. Approval-seeking	.10	ns	-.30	ns
4. Aggressive-hostile behavior	-.31	.05	-.85	.01
5. Attempts to direct & influence ..	-.23	ns	-.65	.01
6. Readiness to withdraw	.02	ns	.58	.01
7. Independent achievement efforts	-.25	ns	.33	.05

* Emotional support-seeking from peers could not be rated as a continuous variable. The observers could, however, report occurrence or nonoccurrence of this behavior. This was compared with compliance after the latter variable was dichotomized at the median. Fisher's exact test was run. High compliance was significantly associated with the occurrence of emotional support-seeking.

immediate right of the social behavior variables contains correlations of these variables with compliance when the children were interacting with other children; the far right column contains parallel data when the children were interacting with adults in the nursery school.⁴

Variables 1, 2, and 3 in Table 1 are dimensions of social behavior which are generally considered to be facets of dependency. Social compliance, too, is usually assumed to be a part of, or associated with, a dependent orientation toward other people. The data provide some support for this assumption. While approval-seeking is unrelated to compliance, the compliant Nursery School children do more readily seek both instrumental help and emotional support from their peers than do the less-compliant children.

⁴ Interrater reliability coefficients for the Nursery School children's ratings were .69 for peer-compliance and .85 for adult-compliance. The median interrater reliability for the seven other social behavior variables was .70 for the peer-interaction ratings and .73 for the ratings of the children's interactions with adults.

Variables 4 and 5, dealing with aggressive and dominating social behaviors, are negatively associated with compliance. However, in neither case is the magnitude of the relationship extremely large. Finally, peer-compliance in nursery school is essentially unrelated to withdrawal tendencies and to independent achievement efforts.

The second column of Table 1 contains correlations representing the Nursery School children's behavior with *adults*, and presents a very different picture from that found in peer interactions. First, the children's compliance toward adults is not related to the amount of dependent overtures the children make toward these adults. Why should the children's compliance be related to amount of dependent behavior with peers but not adults? One possible explanation, admittedly *post hoc*, is that the amount of dependent behaviors exhibited by these children is influenced by the general social context in which they occur. The Fels Institute Nursery School teachers attempt to be unobtrusive, and to leave the children on their own as much as possible. Compliant children with high dependency needs, in this context, might come to expect that dependent overtures would not be readily gratified by these teachers, and might therefore not attempt as many overtures as they would if the teachers were more available.

Second, as seen in the correlations of compliance with variables 4 and 5 in Table 1, strong inverse relationships exist between compliance and hostility and between compliance and dominance when the children are interacting with adults in the nursery school. The Nursery School children who readily comply to adult commands and suggestions frequently do not, or cannot, express much aggression toward these adults. Nor do they usually attempt to get these adults to comply with their own wishes. Conversely, the less-compliant children more readily express aggression toward the nursery school teachers, and more frequently attempt to dominate or influence them. These correlations suggest that, at nursery school age, children's compliance toward adults is a more integral part of a general dominance-submission picture than is their compliance toward peers.

Finally, the Nursery School children's readiness to comply with adults is positively associated with their readiness to withdraw from threatening situations and with the frequency of their achievement efforts. Thus, it appears that the children who are more compliant with adults lack assertiveness, not only toward these adults, but, in general, withdraw more readily in the face of any threat or adversity. The fact that these children also engage in more independent achievement efforts is, at first glance, surprising. Being less assertive children, they might be expected to engage in fewer independent achievement efforts. It is apparent that they were not involved in achievement activities to gain approval from teachers; the correlation of compliance and approval-seeking is negative. There is one plausible alternative explanation. The Nursery School children's readiness to withdraw was correlated .48 with the amount of their achievement efforts. In respect to this, the Nursery School observers informally reported

CHILD DEVELOPMENT

that a number of children, when they became anxious or threatened, withdrew from social interactions of the nursery school and intensely engaged in some achievement activity by themselves. This achievement activity in alone-play following threat seemed to be a method used by these children to allay the anxiety engendered in the previous social situation. The adult-compliant children may have especially used achievement efforts in this defensive fashion.

TABLE 2

Q-SORT ITEMS OF CHARACTERISTIC SOCIAL BEHAVIORS DISCRIMINATING PEER-COMPLIANT FROM PEER-NONCOMPLIANT NURSERY SCHOOL CHILDREN

Item	<i>p</i>
1. Opinions are more readily influenced by others	.05
2. Has a higher energy level	.05
3. Is more spontaneous and uninhibited	.01
4. Is more distractible	.05
5. Is more suggestible	.05
6. More often seeks attention and praise from others	.01
7. Is warmer, friendlier	.05
8. Shows more empathic sensitivity to others' feelings	.05
9. Appears more relaxed, easy-going	.05
10. Is less rigid, inflexible	.05
11. Is less of a perfectionist	.05
12. Exhibits less self pity	.05
13. Finds it less difficult to admit mistakes	.05

The Q-sorts of the Nursery School children's behavior provide additional information regarding the personality attributes of compliant and less-compliant children at this age. Table 2 presents Q-sort items which were found to differentiate the compliant from the less-compliant Nursery School children as they interacted with their *peers*. The items are worded to describe behaviors which were more characteristic of the compliant than the less-compliant children. The numbers in parentheses following the items give statistical levels of significance at which these items discriminated. These were tested with Fisher's exact test using a two-tailed test of significance. The items in Table 2 suggest that the peer-compliant children are somewhat more "outer-directed," to use Riesman's term; they are more suggestible and readily influenced by other persons. And they seek more attention and praise from others. Surprisingly, they have higher energy levels, and are more spontaneous and less inhibited than the less compliant children. These characteristics contrast markedly to the anxious, constricted behavior frequently observed in overly-compliant adults. The Nursery School children who comply readily with their peers are also more friendly,

outgoing, and socially empathic. And they are relatively more comfortable with themselves and others; they are more relaxed, flexible and able to accept their own mistakes. At the nursery school age, then, social compliance with peers does *not* appear to be associated with personal or social inadequacy. If anything, the compliant children present a better picture of mental health than the less compliant ones.

Table 3 presents Q-sort items which were especially characteristic of the Nursery School children who complied readily with *adult* commands and suggestions. The most striking feature of these results, as compared with those of Table 2, is that none of the behaviors accompanying high adult-compliance is similar to those associated with high peer-compliance. Compliance with adults appears to have more of the attributes of social

TABLE 3

Q-SORT ITEMS OF CHARACTERISTIC SOCIAL BEHAVIORS DISCRIMINATING
ADULT-COMPLIANT FROM ADULT-NONCOMPLIANT
NURSERY SCHOOL CHILDREN

<i>Item</i>	<i>p</i>
1. Is more cooperative, eager to please	.02
2. Reacts less negatively to commands from others	.02
3. Is more often a "chronic worrier"	.05
4. Is more deferential to persons considered superior	.02
5. Is less easily irritated by minor frustrations	.02
6. Uses excuses and rationalizations less frequently	.02
7. Becomes upset and anxious more readily	.05

conformity as it is usually pictured. There is an eagerness to please others, a ready acceptance of commands, and a deference toward superiors. The children who comply more readily with adults are also the more anxious children, the "worriers." Why should the personality characteristics of adult-compliant children contrast so markedly to those of the peer-compliant children? One clue may be found in Table 1, variables 4 and 5. It is evident that aggression and dominance are much more negatively related to adult-compliance than to peer-compliance. At nursery school age, compliance to adults appears to have a stronger quality of submissiveness than does compliance with peers. Conversely, noncompliance toward adults is associated with aggressive and dominating behaviors while this is not so in peer interactions. Compliance in peer interactions appears to transpire in a more give-and-take atmosphere. Thus, at nursery school age, peer-compliance and adult-compliance appear to be qualitatively different. Peer-compliance seems to be more a matter of cooperation, while adult-compliance contains a stronger element of submissiveness. This difference makes the differences in personality correlates of adult-compliance vs. peer-

CHILD DEVELOPMENT

compliance more understandable, especially the fact that peer-compliant children present a somewhat better picture of mental health than do the adult-compliant ones.

So far only the social compliance of the Nursery School children has been discussed. Similar data for the older, Day Camp children are presented in Tables 4, 5, and 6. Again, these data have been analyzed separately in respect to peer-compliance and adult-compliance. It will be remembered that the correlation of peer-compliance and adult-compliance at nursery school

TABLE 4
RELATIONSHIPS OF DAY CAMP CHILDREN'S COMPLIANCE
TO OTHER SOCIAL BEHAVIORS

Variable	INTERACTION WITH PEERS		INTERACTION WITH ADULTS	
	Rho	p	Rho	p
1. Instrumental help-seeking	-.30	ns	-.10	ns
2. Emotional support-seeking	*		.10	ns
3. Approval-seeking	-.37	.05	-.30	ns
4. Aggressive-hostile behavior	-.81	.01	-.85	.01
5. Attempts to direct & influence ..	-.71	.01	-.65	.01
6. Readiness to withdraw	.31	ns	.14	ns
7. Independent achievement efforts	-.13	ns	.00	ns

* Instances of emotional support-seeking from peers were so infrequent that this variable could not be rated.

age was .51. For the Day Camp children, the correlation of these behaviors was .76. These older children, then, were more consistently compliant or noncompliant regardless of the age of the persons with whom they interacted. Table 4 summarizes relationships between the Day Camp children's compliance and other aspects of their social behavior.⁵ The correlations in this table, as compared to Table 1, indicate that the Day Camp children's social compliance with *peers* was associated with quite different social behaviors than those found with peer-compliance in the Nursery School children. First, for the Day Camp children, help-seeking no longer relates to high peer-compliance. Second, aggression and dominance are highly related to noncompliance in the behavior of these older children while this was not true in the Nursery School. Day Camp children who readily comply to their peers' commands and suggestions are generally nonassertive and nonaggressive in peer relationships.

⁵ Interrater reliability coefficients for the Day Camp children's ratings were .62 for peer-compliance and .51 for adult-compliance. The median interrater reliability for the other social behavior variables rated was .65 for peer-interaction ratings and .67 for ratings of the children's interactions with adults.

In their interactions with *adults*, the Day Camp children behave differently from the younger children in still other ways. For the older children, compliance with adults is no longer associated with either readiness to withdraw or with independent achievement efforts. High adult-compliance is now exclusively related to low aggression and dominance toward adults. Thus, for the Day Camp children, the dimension of dominance-submission seems to be the central, exclusive one in respect to both peer-compliance and adult-compliance.

TABLE 5

Q-SORT ITEMS OF CHARACTERISTIC SOCIAL BEHAVIORS DISCRIMINATING
PEER-COMPLIANT FROM PEER-NONCOMPLIANT DAY CAMP CHILDREN

<i>Item</i>	<i>p</i>
1. Reacts less negatively to commands from others	.01
2. Opinions are more readily influenced by others	.01
3. Is a less dominating person	.02
4. Expresses views less readily and forcefully	.05
5. Less often attempts to influence others	.05
6. Hates less to lose in a game or argument	.05
7. Takes a more passive role in most relationships	.01
8. Needs to have details "spelled out" more	.05
9. Relies more on others when making decisions	.02
10. Is less impulsive	.01
11. Dramatizes self and own experiences less	.02
12. Uses excuses and rationalizes less	.02
13. Less often projects blame	.05
14. Less frequently criticizes others	.05
15. Is made more uncomfortable by others' hostility	.05

The Q-sorts of the Day Camp children give additional information on the quality of their compliance and noncompliance. Table 5 lists Q-sort items discriminating the children who were highly compliant with their *peers* from those who were less so. Looking at Table 5, high compliance can be seen to be directly reflected in items 1 and 2; the compliant children more readily accede to commands, and their opinions are more easily influenced by others. Items 3 through 7 again indicate that the peer-compliant children are generally passive and nonassertive. Items 8 and 9 suggest that some aspects of dependency are still associated with peer-compliance at this age. Their expression, however, is less direct than the overt help-seeking and support-seeking displayed by the peer-compliant Nursery School children. Items 10 and 11 indicate that, at this older age, some constriction is associated with compliant behavior. Finally, the remaining items suggest that high peer-compliance appears in children who express aggression

CHILD DEVELOPMENT

toward others less readily, and are more uncomfortable with others' aggression.

Data on the personality correlates of the children who complied readily to *adults* in Day Camp are found in Table 6. When the data in this table are compared with those of the previous table, it is evident that adult-compliance is not associated with as many nonassertive behaviors as is peer-

TABLE 6

Q-SORT ITEMS OF CHARACTERISTIC SOCIAL BEHAVIORS DISCRIMINATING ADULT-COMPLIANT FROM ADULT-NONCOMPLIANT DAY CAMP CHILDREN

Item	p
1. Reacts less negatively to commands from others	.01
2. Is more cooperative; eager to please	.01
3. Is warmer, friendlier	.05
4. Is less prone to be dissatisfied	.01
5. Is less dominating	.02
6. Can more readily tolerate opposition to own views	.02
7. Expresses self pity less frequently	.05
8. Less often projects blame	.01
9. Less frequently criticizes those in authority	.05
10. Less frequently criticizes others in general	.02
11. Is made more uncomfortable by others' hostility	.05

compliance. In their stead, there is evidence of an active desire to please others (item 2), and a ready expression of warmth toward others (item 3). Also, neither constriction nor dependency is evident. In general, the adult-compliant children at day camp age appear somewhat healthier than the peer-compliant children. The most general factor found in both the adult-compliant and peer-compliant older children is their low aggressiveness and their uncomfortableness with others' hostility.

The data presented so far have dealt exclusively with the children's social compliance as this was exhibited in the social contexts of the Fels Nursery School and the Fels Day Camp. Presumably these behaviors should be somewhat similar to those displayed by the children in their own homes. This assumption is based on the common observation that children first learn most social behaviors in interaction with significant others in the home, especially parents, and often subsequently carry these behaviors into extra-home situations. The degree that this carry-over may or may not occur for various areas of children's social behavior is, however, an empirical question. The present study provides information on the generality-specificity of young children's compliant behaviors. It will be remembered that the children's degree of compliance with their mothers in their homes was rated by the Fels Home Visitor, while comparable ratings of

the children's compliance with adults and with peers at the Fels Institute were made by other observers. Relationships between these ratings are presented in Table 7. At Nursery School age, the children's compliance with their mothers at home correlates significantly with their compliance toward adults at the Fels Institute. Their compliance with their mothers, however, does not predict peer-compliance. By Day Camp age, however,

TABLE 7

RELATIONSHIPS OF CHILDREN'S COMPLIANCE TOWARD MOTHERS IN THE HOME AND THEIR SOCIAL COMPLIANCE AT THE FELS INSTITUTE

<i>Group</i>	<i>Persons Interacted with at Fels Institute</i>	<i>Rho</i>	<i>p</i>
1. Nursery School ..	Peers	.21	ns
2. Nursery School ..	Adults	.37	.05
3. Day Camp	Peers	.51	.01
4. Day Camp	Adults	.60	.01

the magnitude of the correlations have increased for both peer-compliance and adult compliance. Now the children's compliance toward their mothers in the home is significantly related to *both* their peer-compliance and adult-compliance at the Fels Institute. These latter correlations, when coupled with the finding reported earlier that the relationship between peer- and adult-compliance at Fels was higher for the older children, suggest that social compliance, as a generalized and consistent way of interacting with others, emerges relatively late in young children's social development. Social compliance or noncompliance apparently develops during the preschool age, but only emerges as a consistent personality characteristic when children have entered the larger social world of grade school, teachers, and peers.

These findings logically lead to the general question of how social compliance or noncompliance develops. What produces the increased consistency of the children's social compliance or noncompliance across situations and people as children get older? Obviously compliance, like other social behaviors, must be a product of children's social learning experiences. Significant others, especially parents, presumably affect this development through the process of identification and/or by direct rewards and punishments. The present study provides data relevant to this question, but unfortunately, included neither fathers nor siblings as significant others. Nor does the study provide information on the development of compliance as a function of identification. These problems await additional future research. However, data were obtained regarding the mothers as social reinforcers of their children's compliance. Before turning to the last table summarizing these data, one additional question should be asked: To

CHILD DEVELOPMENT

what degree did the mothers who rewarded their children's compliance also punish noncompliance? The correlation between these two variables for the mothers of the Nursery School children was .05, obviously nonsignificant. Mothers of these young children who readily rewarded compliance were no more or less prone to punish noncompliance than were mothers who seldom reacted positively to their children's compliance. Mothers of the Day Camp children, however, behaved quite differently. For these mothers, rewarding compliance and punishing noncompliance correlated .57, significant at the .01 level of confidence. Thus, the mothers of these older children who rewarded compliance tended to punish noncompliance, while the mothers who were unresponsive to their children's compliance also tended to ignore their noncompliant behaviors. It appears, then, that only the mothers of the older children took a consistent stand on compliance as an aspect of social conformity. This is particularly interesting in view of the fact that individual differences in the children's compliant behavior became much more consistent across situations and people at this later age.

TABLE 8
MOTHERS' REACTIONS TO THEIR CHILDREN'S COMPLIANCE OR
NONCOMPLIANCE IN THE HOME AND THE OCCURRENCE
OF COMPLIANCE OUTSIDE THE HOME

<i>Maternal Reactions</i>	<i>Object of Compliance at Fels Institute</i>	<i>Rho</i>	<i>p</i>
<i>Nursery School Children</i>			
1. Rewarding compliance	Peers	.24	ns
2. Punishing noncompliance	Peers	.20	ns
3. Rewarding compliance	Adults	.27	ns
4. Punishing noncompliance	Adults	.19	ns
<i>Day Camp Children</i>			
1. Rewarding compliance	Peers	.45	.01
2. Punishing noncompliance	Peers	.40	.05
3. Rewarding compliance	Adults	.54	.01
4. Punishing noncompliance	Adults	.27	ns

Table 8 presents relationships between the mothers' reactions to their children's compliance and noncompliance in the home and the degree of social compliance exhibited by these children at the Fels Institute. Again, the data have been analyzed separately for the Nursery School children and for the Day Camp children, as well as with respect to peer- and adult-interactions at Fels. Two trends are evident: First, while the differences are small, in every instance, the mothers' rewarding of compliance predicts their children's behavior better than the punishing of noncompliance. This trend is a provocative one if substantiated in future research. Reward ap-

appears to be more effective than punishment as a technique of parental influence in the development of children's social compliance. This finding is in accord with the reports by Sears and his colleagues (4) that maternal rewards, as a socialization technique, were more effective than punishment for changing young children's dependent and aggressive behaviors.

Second, there are clear-cut differences between the correlations for the Nursery School and Day Camp children. At the earlier ages, neither maternal rewards for compliance nor punishments for noncompliance are significantly associated with the degree that these younger children exhibit social compliance outside the home; maternal reinforcements do not predict these children's social compliance. In contrast, the mothers' reactions to the older children predict the amount of social compliance of these children outside the home in three or four correlations. By early grade school age, mothers' socialization attempts do seem to affect their children's general social compliance; before that they do not—or at least there is no evidence in the present study that they do.

Why should maternal reactions to children's compliance and noncompliance predict better at the older ages? The data provide no information; we can only speculate. One possibility rests on the fact that societal and parental expectations change for children as they grow older. For example, hostility, egocentric behavior, and noncompliance are accepted in our society—at least to some degree—as a natural part of young children's social behavior. However, as children mature and leave mother and home for the broader social experiences of school, they are expected to behave as more socialized, cooperative organisms. These general societal expectations for children's behavior at various ages obviously have their impact on parents. To the degree that parents feel their children to be extensions of themselves, and to the degree that the parents are sensitive to social pressures, they may develop increased concern with their children's social behavior as the children grow older. Parents who themselves value social compliance, either positively or out of a fear of rejection, might now begin to more strongly reward their children's compliant behaviors and punish their noncompliant activities. Other parents, for whom social conformity is less important, will be less concerned with this aspect of their children's social development, and will probably less frequently either reward compliance or punish noncompliance. In other words, differential parental concern may well determine parents' immediate reactions to their children's behavior. These parental reactions, in turn, are major sources of social reinforcements for the children and should, in time, influence the development of the children's social compliance. Parental rewards and punishments obviously affect children's immediate face-to-face compliance with their parents' commands and suggestions. In addition, in time, these parental reactions should also influence the children's social compliance when the parents are no longer present as potential reinforcers. This may transpire because parental reactions not only provide children with cues upon which to base their expecta-

CHILD DEVELOPMENT

tions of receiving rewards or punishments from their parents, but also because such parental reactions establish more generalized expectations in children. Parents are, after all, major transmitters and interpreters of social values to their children. Children, taking cues from parental reactions, develop and incorporate attitudes and expectations concerning the value of various social behaviors, including social compliance. These expectations can, in time, play an important part in determining the quality and degree of social compliance which the children exhibit outside the home when parents are no longer present. As the children's attitudes, values, and expectations become internalized, they may ultimately influence the children's proclivity toward social compliance or noncompliance as did original parental rewards and punishments for social compliance in previous situations. As internalization transpires, the children should exhibit more generalized and consistent compliance or noncompliance across persons and situations than was evident before this social learning occurred. In other words, through internalization, children's social compliance should become less situationally-determined and more the product of the child himself. The present study did find social compliance and noncompliance to be more generalized and consistent in older children. However, our speculations on the role of internalization in this process are *post hoc* ones and must await the test of future research.

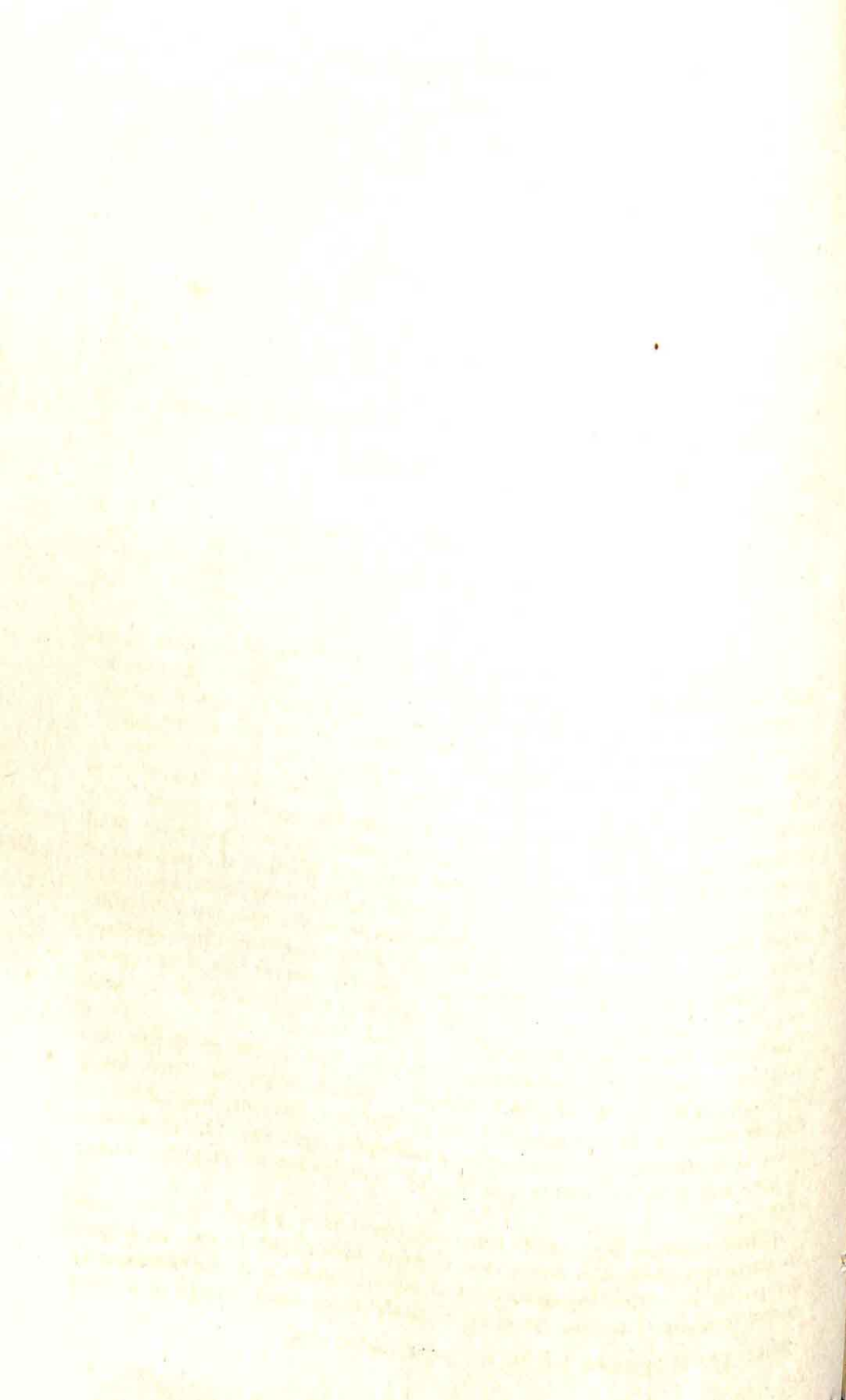
SUMMARY

The present study was concerned with the development of individual differences in young children's social compliance. Two groups of children were studied, a nursery school-age group and a group of six-, seven-, and eight-year-olds. Observers rated and Q-sorted the children's free-play social behaviors as these occurred for the younger group in a nursery school setting and, for the older children, in a day camp. Another observer rated the children's behavior with their mothers in their homes, and the mothers' reactions to this behavior. Relationships among these data were analyzed. These general results were obtained: (a) Sex and intelligence was unrelated to degree of social compliance. (b) Quality of social compliance changed with age. Peer-compliance at nursery school age transpired in a give-and-take atmosphere while peer-compliance in the older children was strongly associated with a general tendency toward low aggression and dominance. (c) The degree of compliance of the older children was more generalized and consistent across situations (from home to Day Camp) and across people (peers vs. adults) than was that of the nursery school age children. (d) In general, maternal rewards for compliance predicted the children's social compliance outside the home better than did maternal punishments for noncompliance. (e) The reactions of the mothers of the older children, both in respect to rewarding compliance and punishing noncompliance, predicted their children's extra-home social compliance to a much greater

degree than did the reactions of the mothers of the nursery school age children. (f) Degree of reward for compliance and punishment for non-compliance were positively and significantly related in the mothers of the older children while no relationship was found between these reactions in the mothers of the younger children. Implications of these results were discussed.

REFERENCES

1. BALDWIN, A., KALHORN, JOAN, & BREESE, FAY. Patterns of parent behavior. *Psychol. Monogr.*, 1945, 58, No. 3 (Whole No. 268).
2. BALDWIN, A., KALHORN, JOAN, & BREESE, FAY. The appraisal of parent behavior. *Psychol. Monogr.*, 1949, 63, No. 4 (Whole No. 299).
3. CRANDALL, V., & PRESTON, ANNE. Patterns and levels of maternal behavior. *Child Develpm.*, 1955, 26, 270-277.
4. SEARS, R., MACCOBY, ELEANOR E., & LEVIN, H. *Patterns of child rearing*. Evanston, Ill.: Row Peterson, 1957.



VISUAL PERCEPTION OF STUTTERERS AND NONSTUTTERERS

BEULAH B. FRANKS *and* CLYDE L. ROUSEY

University of Kansas Medical Center

Although volumes of material have been produced in an effort to explain the phenomenon of stuttering, some areas have failed to attract serious investigation. One such area is that dealing with visual perceptual abilities.

As early as 1915, Swift (5) lectured on what he felt was a disorganization inherent in the visual thinking of stutterers. He initially asked 20 stutterers and 20 nonstutterers the question, "Where do you live?" He followed this by asking them to say "The dog ran across the street." The subjects were then asked to state whether there was any picture evoked in their minds while doing the above tasks and whether the picture was detailed, intense or weak. He noted that all of his normal individuals visualized clearly before and during speech. His stutterers had noticeable difficulty. Only seven reported visualization to his first question. Five visualized well while saying his second statement. All but one of the latter group were under treatment and all of the former group were receiving help. The rest of the subjects had difficulty in visualizing.

In an expanded study following this pilot study, Swift concluded that stuttering severity and ability to visualize are negatively correlated. Swift hypothesized on the basis of his results that this difficulty was caused by some disturbance in the cuneus or visualization area (2). His recommendation was to stimulate the function of picturization in stutterers during therapy.

More recently Bleumel (1) has also written of a thought disturbance in stutterers. Although he assigns primary importance to the breakdown occurring in verbal thought, he also believes there is a blocking during silent thought. This may produce a paralytic-like visual image of a word

CHILD DEVELOPMENT

resulting in an inability to make sounds. He advises exercises in silent thinking and verbal thinking during therapy. Although he has not ascribed the primary reason to faulty visual imagery, there is some indication of this being somehow involved.

The inferences of the above authors were evaluated in this study. A possible disturbance in visual thinking of stutterers was tested by a study of the closure and figure-background abilities of a matched group of stutterers and nonstutterers.

PROCEDURE

Subjects

Twenty stutterers and 20 nonstutterers were matched on the basis of sex, age, race, mental age, intelligence, vision, and hearing. The mean age of both the stutterers and nonstutterers was nine years, four months. There were 16 boys and four girls in both the experimental and control groups. Each member of the experimental group was matched with his control within six months of mental age. Matching on the basis of IQ was such that no subject differed from his control by more than five points.

Selection of the stutterers was done through a survey in which the speech of every pupil in one school district was observed in a classroom situation. Forty children out of approximately 4,300 were diagnosed as stutterers. Only 20 of the stutterers could be suitably matched on the basis of the previously stated criteria with normal children who did not stutter.

Materials

A Keystone Tachistoscope was used for testing visual perception. Seventeen pictures were shown on glass slides at one-fifth of a second and again at one-half of a second. Eight of the slides were pictures used for testing figure-ground responses in a study described by Warner and Strauss (6). The responses for the figure-ground pictures were classified into six categories: Correct Figure, Correct Figure-Ground, Wrong Figure, Wrong Figure-Ground, Ground, No Response. The responses for the closure pictures were placed into two scoring categories: Correct Closure and Wrong Closure.

It was also considered important to study the case histories of the samples. Items used in the case history were taken for the most part from a form suggested by Myklebust (3). A practicing obstetrician designated items indicative of possible brain damage. These included traumatic or congenital physiological instances during prenatal, paranatal or postnatal periods. In addition, a psychiatrist noted items on the case history suggestive of behavioral and/or personality problems. Items included were past familial instances of emotional disturbances as well as behavioral manifestations commonly associated with emotional disturbances. Tabulation of each of these signs was then made for each group.

RESULTS

No fundamental differences in figure-background and closure abilities were found: none of the t values reached the 5 per cent level of confidence. These negative findings hold regardless of the exposure time, i.e., whether one-fifth or one-half second. The results fail to substantiate any inference that a disorganization exists in stutterers similar to that found in known brain-damaged cases. Further, inasmuch as such a disorganization is symptomatic of a visual disturbance in the thought process, the results fail to substantiate either Swift or Bluemel's earlier notions. It of course might be argued that this study in reality measured only the receptive capacity for visual perception of the two groups.¹ If this contention is true, then it would be important to study the visual expression or inner visual thought of stutterers.

Of further interest was the possibility that there might exist between the two groups some fundamental differences in terms of signs of brain damage or emotional disturbance. The two groups were compared using the sign test. The difference between the groups in neither area was statistically significant. The failure to demonstrate differences in these areas could, of course, be due to the fact that the present case history is not a valid instrument for noting subtle but important differences.

SUMMARY

No important differences were noted between nonstutterers and stutterers in a study of their visual perceptual abilities. The two groups were essentially the same with respect to their case histories. Swift's original hypothesis of a negative relationship between stuttering severity and ability to visualize was not supported.

REFERENCES

1. BLEUMEL, C. *The riddle of stuttering*. Danville, Ill.: Interstate, 1957.
2. HAHN, E. F. *Stuttering: significant theories and therapies*. Stanford: Stanford Univer. Press, 1956.
3. MYKLEBUST, H. *Auditory disorders in children*. New York: Grune & Stratton, 1954.
4. STREET, R. A Gestalt Completion Test. *Teach. Coll. Contr. Educ.*, 1931. No. 481.
5. SWIFT, W. B. A psychological analysis of stuttering. *J. abnorm. Psychol.*, 1915, 10, 225-235.
6. WARNER, H., & STRAUSS, A. Pathology of figure-background relation in the child. *J. abnorm. soc. Psychol.*, 1941, 36, 236-248.

¹ Suggested by Dr. Richard Schiefelbusch, University of Kansas.

CHILD DEVELOPMENT

Vol. 29

December, 1958

No. 4

Editor

WILLIAM E. MARTIN
Purdue University

Consulting Editors

BETTYE McDONALD CALDWELL
Washington University School of Medicine

ERCEL S. EPPRIGHT
Iowa State College

EUGENE S. GOLLIN
Queens College

HARRY LEVIN
Cornell University

PAUL H. MUSSEN
University of California

KATHERINE E. ROBERTS
Merrill-Palmer School

WATARU W. SUTOW
University of Texas

WILLIAM W. WATTENBERG
Wayne State University

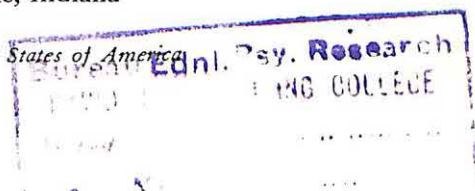
Managing Editor

KATE HOFFMAN

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS
(of The Society for Research in Child Development, Inc.)
Purdue University
Lafayette, Indiana

Made in United States of America



Notice to Contributors

Child Development is devoted to basic research and theory concerning the growth and development of the child.

Manuscripts should be submitted in duplicate, each copy bearing only the title as a means of identification. The title, author's name, and institutional affiliation (together with any footnotes which reveal the identification of the author) should be submitted on separate pages.

Contributors are asked to refer to the *Publication Manual of the American Psychological Association* (1957 Revision) in the preparation of manuscripts.

The following rules require special emphasis:

1. All copy, including indented matter, should be typed double-spaced on white standard paper. Lines should not exceed six inches in length.
2. Footnotes should be few in number. They should be indicated in the text by superscript arabic numerals, numbered consecutively throughout the manuscript. The footnotes themselves should be typed double-spaced on a separate page placed at the end of the manuscript.
3. Tables and figures are costly to print and should be used only when essential. The author will be asked to contribute to the cost of composition of tabular and graphic material, which is excessive in amount or unduly complicated in form. Each table should be typed on a separate sheet. Insert a guide line (e.g., Table 1 about here) at the appropriate place in the manuscript. Figures should be prepared with black India ink, suitable for photography. Legends typewritten on an ordinary typewriter are unacceptable.
4. References in the text should be indicated by numbers in parentheses that refer to the list of references at the end of the article. Do not use footnotes for this purpose. In the list of references, the arrangement should be alphabetical by author's surname, or for multiple authors by the name of the senior; for any one author, by order of date. The following standard form should be observed:

Jones, H. E. Development in adolescence. New York: D. Appleton-Century, 1943.

Shepard, Winifred O., & Schaeffer, M. The effect of concept knowledge on discrimination learning. Child Development, 1956, 27, 173-178.

The World List of abbreviations is used for journals.

The manuscript submitted should represent the final form in which the author wishes the paper to appear. All changes in proof (other than printer's errors) will be charged to the author.

All manuscripts should be addressed to William E. Martin, Editor, Child Development Publications, Purdue University, Lafayette, Indiana.

ADULT AGGRESSION IN THE MANAGEMENT OF DISTURBED CHILDREN¹

ELTON B. McNEIL
University of Michigan

and J. ROBERT COHLER, JR.
Harvard University

In a situation where care must be provided for aggressive children it is imperative to select workers whose control of their own aggressive impulses is adequate to the task. Essentially, a judgment of the worker's potential behavior *in an aggressive situation* is called for and most institutions must either make this selection on the basis of unreliable hunches or rely on a trial period on the job to test the mettle of employees or trainees. The demonstration of a predictable relationship between test measures of personal hostility and ability to manage hostile charges effectively would be of considerable value in preselection of persons for this kind of demanding work.

There have been a number of explorations of the elusive relationship between fantasy and actual behavior. Studies (1, 2, 3, 4, 5, 6, 10, 11, 12, 13) of the connection between hostile fantasies and aggressive actions have dealt with populations ranging from delinquent boys to psychotic adults and have employed a number of measures, such as the TAT, doll play, and story completions, to assess the psychic activity of the individual. Without discussing the details of the findings in each study, it seems clear that the assessment of actual aggressive behavior on the part of the subject has been a chronic problem. Doll play, for example, can simulate a real life setting, but it remains, at best, a miniature and highly symbolic instance of free emotional expression. The fact that the doll victims of aggressive outbursts are inanimate and incapable of responding in kind places a severe limitation on the generalizability of these findings to other more complex

¹ This research was supported by a grant from the Social Science Research Council.

settings. The next most frequent method of determining how much aggression subjects normally express has been by means of tests of reputation. Ratings and rankings of past behavior are gathered from knowledgeable observers and used as criterion measures. In many instances, the reliability of the raters is an unknown quantity, and there has been no opportunity for raters to observe all of the subjects responding to a similar set of emotional pressures. Without some standardization of the situation in which ratings of aggression are made, the meaning of high or low aggression becomes equivocal. Methodological difficulties of this sort have prevented researchers from making a definitive statement regarding the relationship of aggressive fantasy to actual behavior.

For our purposes, the most important theoretical limitation of such studies is that their findings apply particularly to those everyday human relations in which open and violent aggression seldom occurs. Often, it is important to be able to predict human behavior when faced with the stress of undiluted resistance, defiance, or attack. The relationship of fantasy to behavior, under such circumstances, has yet to be fully explored. While the degree of hostility present in a person's fantasy life may not be predictive of loss of self-control in the course of normal human affairs, it might well be an accurate index to his response under substantial interpersonal duress.

The present study is an attempt to explore the relationship between two test measures of hostility and the ability to manage personal aggression in a therapeutic camp for disturbed children. A general hypothesis to be tested is that success in the management of aggressive children bears a predictable relationship to the degree of hostility present in the fantasy productions of the adult child care worker.

METHOD

Research Setting

The University of Michigan Fresh Air Camp provides an ideal setting for field observations of hostility and conflict. Under the guidance of a trained clinical staff and 40 student counselors, the Fresh Air Camp conducts "two, month-long terms of mental-hygiene oriented camping for 120 seriously maladjusted boys, ranging in age from 7 to 14 and referred for this summer experience by one of some 55 treatment agencies" (8). The majority of these children act out their conflicts and openly express hostility toward peers and adults. Throughout the day the counselor is confronted with a series of aggressive incidents requiring intervention and therapeutic handling. The counselor's ability to control his own hostility, as well as the hostility of others, is being tested constantly.

The subjects in this study were 18 male and 15 female counselors who participated in the 1955 session of the Fresh Air Camp. The mean age of the male counselors was 28 while that of the female counselors was 23.

Measures of Hostility

During an orientation and training period at the beginning of the camp session, two measures were used to provide a means of ordering the counselors along a dimension of personal hostility and aggression. The first measure, a test constructed by Watson, Pritzker, and Madison (13), consists of 60 sets of scrambled words which can be assembled to have either a hostile or neutral meaning. For example, the set of words, "Shoot I'll you ask," may be assembled as "I'll ask you" or "I'll shoot you" by omitting either the neutral or the hostile word. The 60 sets of words represent a wide range, in degree of intensity, from very mild to very strong hostility. Watson, Pritzker, and Madison found that these scrambled sentences could be used successfully to differentiate between a neurotic patient group and a group of control subjects with the neurotics showing significantly more hostile constructions than the normals.

A second measure of hostility was derived from the scoring of spontaneous stories elicited in response to the presentation of four TAT (9) cards (3 BM, 8 BM, 18 BM, and 18 GF). A study by Holzberg, Bursten, and Santiccioli (4) offers experimental support for the proposition that individuals whose stories contain an unusual response (greater or less than average) to the aggressive implications of situations suggested by these TAT cards are more highly charged with aggressive tension than are those who respond with an average number of aggressive stories. In their study, aggressive tension was measured in terms of the relative efficiency of learning aggressive, as compared to neutral, material.

Counselor Sociometric Ratings

Working together in the psychologically intimate setting of a therapeutic camp, the counselors soon make judgments regarding one another's personal characteristics and therapeutic skill. These peer judgments often are merciless in their accuracy, for they reflect two months of daily observation of one another's actions in a variety of circumstances. At the conclusion of the camp session each counselor rates the other counselors along a series of dimensions ranging from choice as a personal friend to potential as an understanding parent. For purposes of this study three questionnaire items were chosen as most relevant in estimating counselor reputation for managing aggressive and disturbed children.² For each item the rater was asked to indicate the three counselors who most displayed the characteristics as well as the three who least displayed them. After summing the frequency with which each counselor was nominated as possessing the

² ITEM 3: If I were director of the Fresh Air Camp and given the opportunity of increasing the staff, I'd prefer (3 most) (3 least). ITEM 6: If I were to choose the outstanding counselors, from the standpoint of being understanding and effective with the children, I'd prefer (3 most) (3 least). ITEM 9: The counselor who is best controlled with his children, doesn't lose his temper easily or employ coercive and aggressive methods in dealing with discipline problems is (3 most) (3 least).

most or least of each characteristic for all three items, the authors converted the scores to percentages and rank-ordered all counselors. The rank-order distribution was divided into three equal groups of persons designated as Most Chosen, Average Chosen, and Least Chosen on this measure of counselor preference.

Counselor Behavior

In dealing with disturbed children, everyday situations such as eating, sleeping, or playing are always potential settings for aggressive explosions. Under the guise of interest in determining the causes of such outbursts among the children, we enlisted the aid of 15 counselors who permitted us to observe their interactions with their cabin group. During the camp session observers recorded 10 aggressive incidents from the beginning to end for each of 15 volunteer counselors. The descriptive accounts of these incidents were coded along two dimensions: (a) the effectiveness with which each situation was resolved; (b) the degree of acceptance, by the child, of the solution.

RESULTS

Measures of Hostility and Sociometric Choice

In order to test the relationship between our measures of aggression and the counselors' sociometric ratings of one another's ability to manage aggressive child behavior, chi square analyses were made using the Mood Likelihood Ratio Test (7). The range of scores on the Scrambled Sentence Arrangement Test (from 5 to 27 hostile completions) was divided into three groups—High Aggression, Medium Aggression, and Low Aggression—according to the frequency with which the words were assembled to have a hostile meaning. For the four TAT stories the scores varied between no aggressive stories to four. This restricted distribution necessitated the division of responses into two groups, High Aggression and Low Aggression.

The scores on the Counselor Sociometric Choice measure of ability to handle hostility in disturbed children were characterized as Most Chosen, Average Chosen, and Least Chosen. Our Most Chosen counselors were those judged by their fellow workers to possess the greatest facility in managing aggressive youngsters while the Least Chosen were those judged to have had less ability to manage such children successfully. Comparisons of counselor scores on the test measures of hostility and the peer rankings of skill in managing the children are contained in Table 1. The TAT measure of hostility bears a significant relationship to sociometric choice with the Low Aggression subjects the Most Chosen as successful by their peers while the High Aggression scores are to be found among Average or Least Chosen counselors. Although the direction of the cell loadings is similar to the TAT, the Scrambled Sentence Arrangement Test of hos-

TABLE I

THE RELATIONSHIP OF COUNSELOR SOCIOMETRIC CHOICE TO COUNSELOR
AGGRESSION ON THE TAT AND THE SCRAMBLED SENTENCE
ARRANGEMENT TEST *

	S O C I O M E T R I C C H O I C E				
	Most Chosen	Average Chosen	Least Chosen	χ^2	p
<i>TAT</i>					
High Aggression	2	8	5	6.99	<.05
Low Aggression	9	3	6		
<i>Scrambled Sentences</i>					
High Aggression	3	6	2	6.07	<.20
Medium Aggression	3	4	4		
Low Aggression	5	1	5		
<i>Combined TAT and Scrambled Sentences</i>					
High Combined Aggression	1	6	1	14.96	<.01
Medium Combined Aggression	4	5	5		
Low Combined Aggression	6	0	5		

* In this table, as in the others to follow, two-tailed tests of significance were used. Although prediction of direction of cell loading was made for each table, the small *N* dictated the use of the more rigorous statistical model required by a two-tailed rather than a one-tailed test of level of significance.

tility does not display a significant relationship to the Counselor Sociometric Choice measure.

In an attempt to devise a better measure of hostility, the scores on the TAT and the Scrambled Sentence Arrangement Test were combined to produce a single measure. This was accomplished by making a new rank-order distribution of the raw scores on the TAT and Scrambled Sentence Arrangement Test. Merging these two distributions into a single measure of Combined Aggression and dividing this measure into three equal groups of persons, categorized as High, Medium, and Low, required some decision making. At the high and low extremes, for each measure, categorization posed no problem. In those cases in which a counselor scored high on one measure and low on another he was relegated to the Average category, since designation as either high or low on the Combined Aggression measure could not be defended. This procedure allows the extremes to become fairly pure cases while the Average category is composed of persons average on both measures or high on only one. Categories of High, Medium, and Low Combined Aggression were formed and compared with sociometric choice. This combined measure reveals that a High Combined Aggression

score is most regularly associated with the Average Chosen category, while Medium Combined Aggression scores appear equally in the Most, Average, and Least Chosen categories. A Low Aggression rating is distinctly related to being either Most or Least Chosen by others.

Hostility, Sociometric Choice, and Child Management

A clearer understanding of the meaning of the counselors' sociometric choices of one another can be derived from an examination of their relationship to actual counselor behavior with the children. The expert observations of counselor handling of conflicts between children were coded according to the effectiveness with which these conflicts were resolved. Effectiveness, in this instance, was judged in terms of the degree to which peace was restored in the group, the time required to negotiate the cessation of hostilities, the permanence of the resolution, and the need for help in achieving a solution. On the basis of a cumulative score from 10 such incidents, each counselor was classified as High, Medium, or Low in Effectiveness.

In many instances an uneasy peace was established at the expense of the child who had initiated the conflict. Although such instances might reasonably have been included as an indication of ineffective conflict resolution, it was a sufficiently frequent and distinct characteristic of counselor-child relationships to warrant special consideration. The child's perception of his treatment and his behavior following the counselor's attempt at resolution were coded into categories of High, Medium, and Low Acceptance.

Table 2 reveals that both conflict resolution and the child's acceptance of the resolution bear similar and highly significant relationships to sociometric choice. Our Most Chosen counselors regularly are found to be very effective in resolving conflict and have a high acceptance of these resolu-

TABLE 2
THE RELATIONSHIP OF COUNSELOR SOCIOMETRIC CHOICE TO CONFLICT
RESOLUTION AND CHILD ACCEPTANCE OF RESOLUTION

	S O C I O M E T R I C C H O I C E			χ^2	<i>p</i>
	<i>Most Chosen</i>	<i>Average Chosen</i>	<i>Least Chosen</i>		
<i>Conflict Resolution</i>					
High Effectiveness				13.44	< .01
Medium Effectiveness	4	0	1		
Low Effectiveness	0	4	1		
	1	1	3		
<i>Child Acceptance of Resolution</i>					
High Acceptance				10.67	< .05
Medium Acceptance	4	0	1		
Low Acceptance	0	3	2		
	1	2	2		

tions by the child. The reverse of these findings is suggested for our Least Chosen counselors. Thus it seems that the choice of successful and unsuccessful counselors by their peers is an accurate reflection of their handling of difficult situations, as judged by an outside observer.

Sociometric choice, then, relates significantly to the TAT measure of aggression, the combined measure of aggression, the counselor's ability effectively to resolve conflict, and the child's acceptance of the resolution of the conflict. An important next step is to determine the nature of the relationship between our precamp measures of counselor aggression and the later behavior of the counselors in dealing with hostile children.

TABLE 3

THE RELATIONSHIP OF COUNSELOR AGGRESSION ON THE TAT TO
CONFLICT RESOLUTION AND CHILD ACCEPTANCE
OF RESOLUTION

	T A T		χ^2	p
	High Aggression	Low Aggression		
<i>Conflict Resolution</i>				
High Effectiveness	1	4	8.46	< .02
Medium Effectiveness	5	0		
Low Effectiveness	3	2		
<i>Child Acceptance of Resolution</i>				
High Acceptance	1	4	5.18	< .10
Medium Acceptance	4	1		
Low Acceptance	4	1		

Table 3 indicates that a significant relationship exists between the TAT measure of aggression and Effectiveness in Conflict Resolution, but that the association with degree of Acceptance of Conflict Resolution by the child attains only a borderline level of significance. Counselors who were rated Low in aggression on the TAT tended to be more effective in resolving hostility and in having the child feel that he was dealt with reasonably and fairly. In general, counselors rated High in aggression received no more than a Medium or Low rating of effectiveness and found a similar degree of acceptance by the children. In all fairness it ought to be noted that the category designated as High Acceptance of Resolution was quite relative to the group and the setting. Although the degrees of acceptance were distinguishable from one another, their over-all range was somewhat restricted. It was seldom possible to achieve enthusiastic acceptance from our revenge-motivated children, and this fact made the task of prediction a particularly difficult one.

The scrambled sentences measure of aggression, which failed to bear a significant relationship to counselor sociometric choice, was similarly without predictive value when compared with either Conflict Resolution or Child Acceptance of Resolution.

The combined test measure of counselor aggression also failed to differentiate counselors in terms of the effectiveness of their attempts to resolve conflict and in the acceptance of the resolutions by the children.

DISCUSSION

The theoretical and methodological problems of assessing the relationship of aggressive fantasy to hostile behavior are apparent in the results of this study. Table 1 underlines a number of these problems because it contains implications for both the assets and limitations of such efforts. Both the TAT and the Combined TAT and Scrambled Sentences measures are able, empirically, to predict successfully to peer ratings of clinical skill; yet this finding leaves us adrift theoretically. It is clear that the prediction of behavior from fantasy with these tools is useful as a wide screening device but does not provide answers to our fundamental questions regarding the meaning of hostility in the individual's psychic economy. In particular, we have not been able to determine whether a nonaggressive story, told to a TAT picture, reflects low intensity of aggressive needs or intense needs coupled with determined self-control. It would seem that such knowledge would be essential to making intelligent decisions about matching workers to patient populations.

The evidence in Table 1 suggests, as other workers have discovered, that all measures of fantasy "aggression" are by no means equivalent. There exists a "hidden" quality in the aggression expressed in the TAT when compared with that revealed in the Scrambled Sentences test. Hostile thoughts, which fit naturally in the flow of events of a story suggested by a TAT card, evoke a much greater self-consciousness when they appear in the form of a sentence which the subject has constructed to read, "I'll shoot you." The high visibility of such responses seems to act as an inhibitor which moves the subject consciously to alter the degree of hostility he allows to make an appearance. This phenomenon is probably most characteristic of somewhat sophisticated subject groups such as our own. In part, the decision to combine the TAT and Scrambled Sentences test into a single measure of aggression was prompted by the reports of several subjects that they changed the direction of their responses after constructing "too many" aggressive sentences. The combined measure held promise of providing some compensation for the subjects' insertion of conscious control over his responses. This measure proved to be an effective device for predicting to the counselor Sociometric Choice but ineffective in foretelling the counselor's ability to resolve conflict and manage a child's open anger.

We found (Table 3) that only one measure (the TAT) was able to predict later behavior in a hostile situation. As is so frequently reported in the literature, all tests reputed to measure "aggression" measure something—although apparently not the same thing. In part, this occurs as a result of the lack of refinement of the various tests, but our logic becomes circular if we allow the blame to rest there. Perhaps it is unreasonable to expect any simple device adequately to account for the multitude of dimensions implicit in an abstraction such as "aggression." The complexity of a typical expression of hostility often involves most of what we know of human behavior and a great deal that we do not. It should not surprise us, then, that tests assessing a broad spectrum of potential response are likely to have some predictive value. In many respects the TAT, drawing as it does from a considerable range of the individual's fantasy life, satisfies this requirement, yet leaves much to be desired as a dependable tool.

A part of this study provides experimental confirmation of what every clinical worker knows. In the test tube of the working situation one's reputation for effectiveness and skill becomes a fundamental aspect of his relations with fellow workers. In the highly visible setting of a therapeutic camp the "office hour" is exposed to scrutiny, and the judgments of peers and supervisors are soon in close agreement. The significance of the relationships between these judgments of clinical adequacy and the aggressive fantasy of the subjects can in part be accounted for by this "open" quality of the camp setting. The frequency of aggressive outbursts by the children presented a host of opportunities to make accurate judgments of the adequacy of adult aggressive control. In Table 2 it is apparent that the accuracy of counselor judgment of the clinical skill of other counselors can be verified by using independent observations of the handling of crisis situations. One's reputation as an effective clinical worker is not always judged with such ease or accuracy since much of therapeutic work is of a highly private nature and leaves only vague inference as a basis for judging. It seems clear that if co-worker judgments are to be used as a means of assessing ability, a significant portion of the process of clinical work must be made visible to others. Too frequently this is not the case; yet we rely heavily on peer opinions in making our judgments.

In such circumstances, the discrete meaning in any individual worker of aggressive fantasy was less important than some estimate of its gross amount. Most clinical settings demand greater refinement of prediction and might profitably make use of the leads provided by this study.

SUMMARY

Our concern with the role of adult aggression in the management of disturbed children arose from the need to appraise the probable success of potential child care workers. From clinical observations it was clear that

the mixing of aggressive children and hostile adults inevitably produced a flood of therapeutic crises which tended to scatter treatment plans into wild disarray. In contrast, the successful management of hostile situations seemed regularly to be characteristic of those adults whose own aggressive needs were the least pressing or under the best control. To test this prediction, two measures of aggression (a Scrambled Sentences Arrangement Test and spontaneous stories in response to four TAT cards) were administered to a group of counselors prior to their summer experience in a therapeutic camp for disturbed and delinquent children. Comparisons were made between the degree of aggressive response to these tests and ratings, by their fellow counselors, of success in managing hostile situations. An additional measure of counselor skill was gained through observation by experienced observers of 10 crisis incidents for each of a group of volunteer counselors. These detailed observations formed the basis for rating both counselor effectiveness in resolving conflict and the degree to which the children were able to accept and "live with" these resolutions of their disputes.

Of the two test measures used to assess the degree of hostility present in the fantasy productions of our sample of child care workers, it is clear that successful prediction to aggressive behavior occurs more readily using the TAT measure. Significant relationships exist between the degree of hostility revealed in spontaneous TAT stories, child care workers' ratings of one another's clinical skill, and experienced observers' ratings of the subject's handling of aggressive situations. A positive and significant relationship was discovered between expert and peer judgments of an individual's clinical ability. The refinement of available testing techniques and further theoretical clarification of the relation of fantasy to behavior is needed to increase the accuracy with which we can predict worker behavior in an aggressive situation.

REFERENCES

1. BACH, G. R. Young children's play fantasies. *Psychol. Monogr.*, 1945, 59, No. 2 (Whole No. 272).
2. ERIKSEN, C. W. Some implications for TAT interpretations arising from needs and perception experiments. *J. Pers.*, 1951, 19, 281-288.
3. FESHBACH, S. The drive-reducing function of fantasy behavior. *J. abnorm. soc. Psychol.*, 1955, 50, 3-11.
4. HOLZBERG, J. D., BURSTEN, B., & SANTICCIOLI, A. The reporting of aggression as indication of aggressive tension. *J. abnorm. soc. Psychol.*, 1955, 50, 12-18.
5. KAGAN, J. The measurement of overt aggression from fantasy. *J. abnorm. soc. Psychol.*, 1956, 52, 390-393.

6. KORNER, A. F. *Some aspects of hostility in young children*. New York: Grune & Stratton, 1949.
7. MOOD, A. M. *Introduction to the theory of statistics*. New York: McGraw-Hill, 1950.
8. MORSE, W. C. The development of a mental hygiene milieu in a camp program for disturbed boys. *Amer. J. Orthopsychiat.*, 1953, 23, 826-832.
9. MURRAY, H. A. *Thematic Apperception Test: manual*. Cambridge, Mass.: Harvard Univer. Press, 1943.
10. MUSSEN, P. H., & NAYLOR, H. K. The relationship between overt and fantasy aggression. *J. abnorm. soc. Psychol.*, 1954, 49, 235-240.
11. SANFORD, R. N., ADKINS, MARGARET M., MILLER, R. B., COBB, E. A., & OTHERS. Physique, personality and scholarship: a cooperative study of school children. *Monogr. Soc. Res. Child Developm.*, 1943, 8, No. 1.
12. SEARS, R. R. Relation of fantasy aggression to interpersonal aggression. *Child Developm.*, 1950, 21, 5-6.
13. WATSON, R. E., PRITZKER, L., & MADISON, P. Hostility in neurotics and normals. *J. abnorm. soc. Psychol.*, 1955, 50, 36-40.

A SELF-CONCEPT SCALE FOR CHILDREN AND ITS RELATIONSHIP TO THE CHILDREN'S FORM OF THE MANIFEST ANXIETY SCALE¹

LEWIS P. LIPSITT²

Iowa Child Welfare Research Station

The kinds of verbalizations that people "make to themselves," perhaps particularly about themselves, have long been of interest to personality theorists such as Murray, Rogers, Sullivan, and Allport. Such self-talk, as verbal behavior, typically becomes important in personality and clinical appraisals. Self-evaluative behavior, in fact, seems to be involved in the operational definitions of certain personality concepts as ego, superego, ego-ideal, self-concept, ideal-self, and the like. Thus, an individual who verbalizes inadequacy and inferiority is said to have weak ego-strength or a low self-concept. Moreover, persons' verbal representations of themselves are often compared with their concepts of an "ideal self" on the not un-

¹ A portion of this paper was read at the 1957 meetings of the Midwestern Psychological Association. The writer is indebted to Mr. M. M. Schell, Superintendent, and Principals Elsie E. Glassburner and Edward L. Hutchcroft of the Washington, Iowa, Public Schools for their cooperation in providing subjects and facilities. The writer is appreciative, also, of critical readings of an early manuscript by Drs. Ruth B. Holton and Alfred Castaneda. Thanks are due Mr. Andy E. Rogers who aided in the collection and processing of data.

² Now of Department of Psychology, Brown University.

common assumption that in the conflicted or anxious individual there exists a large discrepancy between his ideal and his conception of himself.

There has arisen, however, some evidence contrary to the assumption of a negative relationship between anxiety level and verbal self-estimations. Taylor and Combs (18) presented children with written statements which the authors considered true of all children but which would require self-effacement to admit, finding that those who tended to admit to these statements were those with the highest adjustment scores as measured by the California Test of Personality. Zimmer (19) was unable to find a relationship between a self-acceptance measure and independent measures of conflict derived from a word association task.

Researchers interested in self-concept variables typically utilize discrepancy scores (1, 2, 5, 6, 7, 8, 10, 13, 15, 19) deriving from two test measures, the first involving *S*'s rating himself ("as he really thinks he is") with regard to trait-descriptive adjectives and the second requiring *S* to respond to the same adjectives indicating "how he would like to be." The discrepancy score may consist of the total difference between the two measures, and this discrepancy is considered as reflecting degree of dissatisfaction with oneself.

The present study employs both self-concept and discrepancy measures in fourth, fifth, and sixth grade children and is concerned primarily with (a) the comparative reliabilities of the discrepancy scores and the self-concept scores, and (b) the relationship of each of these two measures to scores on the children's form of the manifest anxiety scale (3, 4, 11, 14).

The self-concept scale contained the following 22 trait-descriptive adjectives, presented here in the order used in the scale: friendly, happy, kind, brave, honest, likable, trusted, good, proud, *lazy*, loyal, co-operative, cheerful, thoughtful, popular, courteous, *jealous*, obedient, polite, *bashful*, clean, helpful. Each of these adjectives was prefaced by the phrase "I am . . ." and was followed by a five-point rating scale. Nineteen were considered as positive or socially desirable attributes, while three were considered negative (italicized above). The rating categories, scored from 1 to 5, were entitled *not at all*, *not very often*, *some of the time*, *most of the time*, and *all of the time*. A score of 1 was received on an item if *S* checked the first category, a score of 5 if the last category were checked, except in the case of the three negative adjectives which were scored in inverse fashion.³ A score on the self-concept scale was obtained for each *S* by summing the ratings ascribed to himself on each item. Lower scores were presumed to reflect degree of self-disparagement.

The ideal-self scale contained the same adjectives but here each was prefaced by "I would like to be . . ." Again, *S* rated himself on each item on similar scales. The discrepancy scores employed in the present

³ The negative words may provide a detection device for nonreaders and indiscriminate checkers. In the present study, however, no *S*s were eliminated on such a basis.

study consisted of the simple subtraction of the total self-concept score from the total ideal-self score.⁴

Both scales were administered, with the children's manifest anxiety scale (CMAS), to approximately 300 fourth, fifth, and sixth graders on the same day, and again two weeks later. Administration was on a group basis in the classrooms, and Ss were permitted as much time as needed to complete all the scales. The CMAS was administered first, the self-concept scale next, and the ideal-self scale last.

RELIABILITY OF THE CHILDREN'S FORM OF THE MANIFEST ANXIETY SCALE

A report of the children's form of the manifest anxiety scale (CMAS), published in this journal (3), contained means, standard deviations, and one-week reliability coefficients for a city population of fourth, fifth, and sixth graders. The present study provides additional information concerning the retest reliability of the CMAS for a two-week interval and a different population of same-aged children, and provides some indication of the generality of the originally reported mean scores for these age groups.

TABLE 1
ANXIETY SCALE MEANS AND STANDARD DEVIATIONS
BY GRADE AND BY SEX

	FOURTH GRADE			FIFTH GRADE			SIXTH GRADE		
	N	M	SD	N	M	SD	N	M	SD
Boys	47	13.13	7.38	50	15.16	7.96	41	14.51	8.12
Girls	62	15.61	8.83	61	15.26	7.44	37	16.57	8.16

Table 1 contains the anxiety means and standard deviations for each grade and sex separately. These statistics may be compared with those describing the original population (3, p. 320). All of the present means fall somewhat below those previously reported. Significant differences between the present and previous means are found for fourth grade girls only ($t = 2.34$, $p = .05$).⁵

⁴ Seven Ss who rated themselves higher on the self-concept than the ideal-self scale were eliminated.

⁵ Despite the general statistical similarity of these two sets of group means, both of which are derived from presumably normal populations, this one disparity suggests caution in the interpretation of differences found between different test populations (e.g., normal school children vs. clinical samples). That is, significant differences should probably not be taken as necessarily reflecting the presence of population abnormality. It is suggested, furthermore, that the obtained mean for fourth grade girls in the original population may be an overestimate for a population of normal fourth grade girls.

CHILD DEVELOPMENT

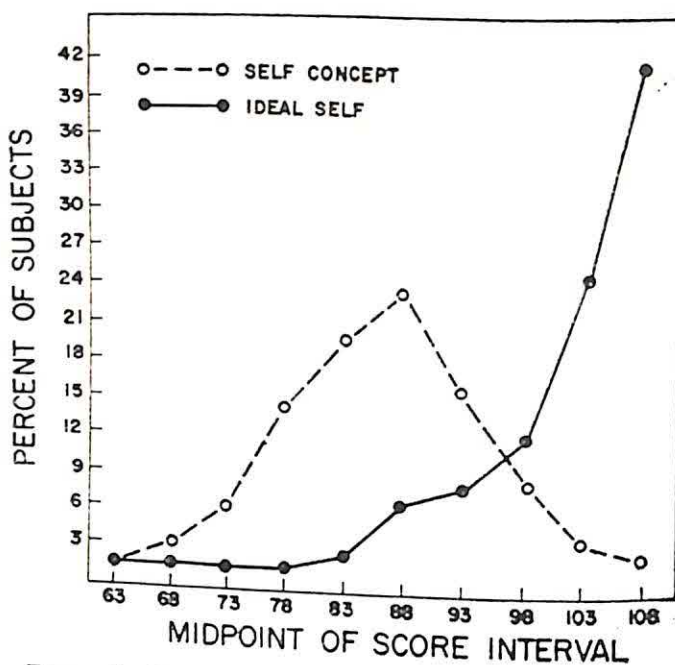


FIGURE 2—Frequency polygon showing percentage of 298 fourth, fifth, and sixth grade children receiving the indicated scores on the self-concept and ideal-self scales.

Our interest here is in the respective reliabilities of the *self-concept* and *discrepancy* scores. Table 5 depicts the two-week test-retest reliability coefficients for the self-concept measure alone. All of these correlations are significant beyond the .001 level.

Also in Table 5 are the comparable retest correlations for the discrepancy measure. These coefficients are all somewhat below those for the self-

TABLE 5
TEST-RETEST CORRELATIONS FOR SELF-CONCEPT SCALE AND
DISCREPANCY SCORES BY GRADE AND BY SEX

	FOURTH GRADE		FIFTH GRADE		SIXTH GRADE	
	N	r	N	r	N	r
<i>Self-Concept Scale</i>						
Boys	44	.73***	46	.80***	40	.84***
Girls	56	.78***	49	.91***	35	.80***
<i>Discrepancy Scores</i>						
Boys	28	.57**	44	.72***	38	.56***
Girls	32	.66***	48	.68***	34	.51***

** Significant at .01 level.

*** Significant at .001 level.

concept measure.⁶ Four of six attain significance, nevertheless, at the .001 level, and two at the .01 level.

RELATIONSHIPS OF SELF-CONCEPT AND IDEAL-SELF MEASURES WITH
SCORES ON THE CHILDREN'S MANIFEST ANXIETY SCALE

Table 6 reveals a negative relationship between the self-concept scores and anxiety scores as measured by the CMAS, for each of the grades and sexes separately. In all cases, these correlations are significant, five at the .01 and one at the .05 level. The magnitude of these correlations may be compared with those relating anxiety to the discrepancy scores, only four of which attain significance, two at the .01 and two at the .05 level. It would appear, then, that the self-concept score by itself yields a measure that is more reliable than the discrepancy score and that is more highly related to the CMAS score.

TABLE 6
CORRELATIONS BETWEEN ANXIETY SCALE AND SELF-CONCEPT SCALE AND
DISCREPANCY SCORES BY GRADE AND BY SEX

	FOURTH GRADE		FIFTH GRADE		SIXTH GRADE	
	N	r	N	r	N	r
<i>Anxiety Scale and Self-Concept Scale</i>						
Boys	47	-.53**	50	-.40**	41	-.34*
Girls	62	-.63**	61	-.40**	37	-.58**
<i>Anxiety Scale and Discrepancy Scores</i>						
Boys	43	.31*	48	.24	40	.20
Girls	55	.51**	60	.39**	36	.39*

* Significant at .05 level. ** Significant at .01 level.

CONCLUSIONS AND DISCUSSION

The conclusions that may be drawn from these data are: (a) both the self-concept and discrepancy measures, as presented here, provided reliable estimates of an individual difference variable; (b) of these two types of measures, the self-concept score provided a somewhat more reliable measure than the discrepancy score; (c) for the fourth, fifth, and sixth grade Ss, there were no reliable grade or sex differences in mean self-concept scores; (d) for all six grade-sex combinations, the self-concept measure correlated significantly with performance on the CMAS, while comparable correla-

⁶ The loss of Ss for the fourth grade in Table 6 was due to the impossibility of administering the ideal-self scale to one of the classes on the second day of testing.

tions for the discrepancy scores were somewhat less and in some cases not reliable.⁷

The tendency toward self-disparagement, as reflected in responses to the self-concept scale, was found to be significantly related to another independent measure of the degree to which a person admits to possessing symptoms commonly held to reflect anxiety.⁸ Several theorists (9, 16) have ascribed drive properties to anxiety; thus, theoretical predictions concerning the role of drive in learning may be applicable here as well.

The tendency toward self-disparagement may be conceived as a phenomenon producing behavioral effects consonant with its status as a drive variable as follows: The conditioning principles presumed to underlie the acquisition of secondary drives assume that neutral stimuli preceding reinforced responses may come through training to elicit those responses themselves in the absence of the original reinforcement. Thus, if a verbal stimulus ("You are naughty") is presented in conjunction with a noxious stimulus (physical punishment) which unconditionally elicits an emotional response (pain), that verbal stimulus may acquire the capacity to elicit (some part of) the emotional response independent of the original noxious stimulus. It is quite conceivable, then, that the conditions (behavior of the child) which originally resulted in the noxious stimulation (punishment) may come to elicit the verbal response and its stimulus counterpart ("I am naughty") which in turn mediates the emotional response. The tendency toward self-disparagement becomes, then, an antecedent condition for general emotional responsiveness (anxiety) or drive.

SUMMARY

The children's form of the manifest anxiety scale and self-concept and ideal-self scales were administered twice to approximately 300 fourth, fifth, and sixth graders at a two-week interval. A discrepancy score, or measure of self-disparagement, was obtained by subtracting Ss' self-concept ratings from their ideal-self ratings. It was found that the self-concept measure taken by itself was more reliable than the discrepancy measure, and that

⁷ The lower retest reliabilities of the discrepancy measure may be attributed to the fact that difference scores always involve additional sources of variability to those involved in a single test. The coefficient for the discrepancy measure is a function of the variability of the two tests, since the discrepancy score is derived from the use of both the self-concept and ideal-self scores. McNemar (12, p. 154) points out that the standard error of such difference scores must be lower than for one test alone, and that this tends to depress the reliability estimates of difference scores.

⁸ Some may consider this relationship as indicative of the presence of a response set or test-taking attitude which mediates or influences scores on both scales. The view taken here is that whenever two test measures correlate, it may be said that a response set is responsible in that there exists a source of variance common to both measures. When test measures are correlated, they are measuring something in common, response set or whatever.

the self-concept measure was more highly related to CMAS score than was the discrepancy score. Significant correlations were obtained for all grades and sex combinations between CMAS and self-concept scores, with high anxious Ss producing low self-concept (or high self-disparagement) ratings. The two-week reliability of the CMAS was essentially the same as that for the original population.

REFERENCES

1. BILLS, R. E., VANCE, E. L., & McLEAN, D. S. An index of adjustment and values. *J. consult. Psychol.*, 1951, 15, 257-261.
2. BLOCK, J., & THOMAS, H. Is satisfaction with self a measure of adjustment? *J. abnorm. soc. Psychol.*, 1955, 51, 254-259.
3. CASTANEDA, A., McCANDLESS, B. R., & PALERMO, D. S. The children's form of the manifest anxiety scale. *Child Developm.*, 1956, 27, 317-326.
4. CASTANEDA, A., PALERMO, D. S., & McCANDLESS, B. R. Complex learning and performance as a function of anxiety in children and task difficulty. *Child Developm.*, 1956, 27, 327-332.
5. CHORDOKOFF, B. Adjustment and the discrepancy between the perceived and ideal self. *J. clin. Psychol.*, 1954, 10, 266-268.
6. COWEN, E. L., HEILIZER, F., & AXELROD, H. S. Self-concept conflict indicators and learning. *J. abnorm. soc. Psychol.*, 1955, 51, 242-245.
7. FRIEDMAN, I. Phenomenal, ideal, and projected conceptions of self. *J. abnorm. soc. Psychol.*, 1955, 51, 611-615.
8. HANLON, T. E., HOFSTAETTER, P. R., & O'CONNOR, J. P. Congruence of self and ideal self in relation to personality adjustment. *J. consult. Psychol.*, 1954, 18, 215-218.
9. HULL, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
10. LEVY, L. H. The meaning and generality of perceived actual-ideal discrepancies. *J. consult. Psychol.*, 1956, 20, 396-398.
11. McCANDLESS, B. R., & CASTANEDA, A. Anxiety in children, school achievement, and intelligence. *Child Developm.*, 1956, 27, 379-382.
12. McNEMAR, Q. *Psychological statistics*. New York: Wiley, 1955.
13. MILLER, K. S., & WORCHEL, P. The effects of need-achievement and self-ideal discrepancy on performance under stress. *J. Pers.*, 1956, 25, 176-190.
14. PALERMO, D. S., CASTANEDA, A., & McCANDLESS, B. R. The relationship of anxiety in children to performance in a complex learning task. *Child Developm.*, 1956, 27, 333-337.

CHILD DEVELOPMENT

15. ROBERTS, G. E. A study of the validity of the index of adjustment and values. *J. consult. Psychol.*, 1952, 16, 302-304.
16. SPENCE, K. W. Current interpretations of learning data and some recent developments in stimulus-response theory. In *Kentucky Symposium: Learning Theory, Personality Theory, and Clinical Research*. New York: Wiley, 1954.
17. SPENCE, K. W. *Behavior theory and conditioning*. New Haven: Yale Univer. Press, 1956.
18. TAYLOR, C., & COMBS, A. W. Self-acceptance and adjustment. *J. consult. Psychol.*, 1952, 16, 89-91.
19. ZIMMER, H. Self-acceptance and its relation to conflict. *J. consult. Psychol.*, 1954, 18, 447-449.

SIZE JUDGMENT AT DIFFERENT DISTANCES AS A FUNCTION OF AGE LEVEL

WALTER COHEN, AARON HERSHKOWITZ, *and*
MARJORIE CHODACK
University of Buffalo

Although many investigators and theorists have dealt with the problem of perceptual constancy, an adequate understanding of the phenomenon is lacking. Much of the controversy has been concerned with the traditional problem of nativism vs. empiricism. Whether or not this issue can be resolved, the development of constancy would seem to merit study in and of itself.

Studies of constancy as a function of age have yielded inconsistent data. Size constancy represents a clear case in point. Various investigators (1, 2, 9) have found that size constancy increases with age while others have presented evidence that there is no such relationship (3, 4). Variations in experimental methodology and measurement techniques have made direct confirmation of theory and comparison of data impossible. The experiments differ considerably in the shape, size, and distance of stimuli, types of comparison made, instructions, and range of ages tested. Any one of these variables might well affect the outcome of a developmental study. In addition, very small samples and inadequate statistical analysis characterize most of the research in this field.

There appears to be an implicit assumption that size constancy is a simple, unitary phenomenon which develops, or does not develop, with age, and there is thus a tendency to ignore the specific experimental conditions under which constancy is measured. The present experiment is designed to investigate some of the variables which might account for the seemingly contradictory data in the literature and to clarify the issue of constancy as a unitary process.

One of the variables that seems most significant is that of the distance of the objects. Lambercier and Piaget (5) found size constancy differences between children and adults for "far" objects. Young children (5 to 7) tended toward underconstancy; older children (10 to 11) showed near perfect constancy, while adults (17 and above) showed definite overcon-

stancy. No differences were found between children and adults for "near" distances, and Piaget contends that under these conditions constancy is complete by the end of the first year. The terms "near" and "far" appear to refer mainly to the distance between the observer and the most distant object. More recently, Zeigler and Leibowitz (9) found that 7- and 8-year-olds show good constancy at "near" distances (10 ft.) while adults show overconstancy at this distance. Both groups display underconstancy at greater distances, but the adults are more accurate in their judgments. In both of these investigations, the variable of distance is ambiguous since both the distance of objects from the observer, as well as the distance between the objects themselves, are simultaneously varied. The present study was designed to investigate the effects of differences in depth between the test objects themselves.

Werner (7) has stressed the diffuse quality of the child's perception. However, this aspect has been neglected in previous studies of size estimation development. Usually the child is restricted to a single estimation of equivalence and only the PSE is computed. The present investigators used the Method of Limits, which permits the *O* to give a number of responses of "same" to the standard during a single series (8). It was felt that a measure of the range of equivalence, i.e., the "interval of uncertainty" (IU), might reveal aspects of size judgments which are relevant for the problem of perceptual development.

METHOD

Apparatus

The experiment was conducted in a well-illuminated room, 9 by 30 ft., in which the walls and windows were entirely covered with white paper. *O* sat on an adjustable chair. His chin was placed on a headrest 100 cm. from the floor. Three black cardboard squares, 5, 6, and 7 cm. each, served as standard stimuli. Only one of these was present at each trial. Two desks, each 73 cm. high, were placed at 2 and 6 m. from the *O* and were visible at all times. A standard was clamped to one of the two desks so that it was at a height of 129 cm. from the floor. The standard was always 30 cm. to the left of *O*'s line of sight.

Thirteen black cardboard comparison stimuli were mounted on white dowels which were in turn mounted on a rotating drum. The drum could be rotated so that only one comparison stimulus was exposed at a time at a height of 169 cm. from the floor. The sizes of the comparison stimuli ranged in $\frac{1}{2}$ cm. steps from 3 cm. to 9 cm. and were arranged consecutively on the drum. A white cloth, which obscured both the *E* and the comparison wheel from *O*'s view, was stretched across the width of the room directly in front of the drum.

To summarize, each *O* was confronted with a single comparison stimulus (at a distance of 8 m. and directly in front of *O*) and a single standard

(at a distance of either 2 or 6 m. and slightly to the left of *O*). Stimuli were seen against a white background. The only other objects in the field were two desks, placed 2 and 6 m., respectively, in front of *O*.

Subjects

Forty-two *O*s were randomly selected from various classes of the Barker, N.Y., school system. There were 12 *O*s each in the 7-, 12-, and 17-year-old groups, and 6 *O*s in the 5-year-old group. An equal number of boys and girls was selected for each age group. Only one *O*, a 5-year-old, was not testable and was replaced.

Procedure

Each *O* was seated at an appropriate height so that he was comfortable when using the head rest. He was instructed to tell whether the card above the white cloth looked larger than, smaller than, or the same size as the card on the desk. The instructions were repeated and the objects were pointed to when necessary. The tester was a 22-year-old woman who was previously known to most of the *O*s.

An ascending and descending series of judgments were obtained from *O* for each of the 3 standards at both distances (2 and 6 m.). The order of presentation was counterbalanced for distance and size of the standard as well as for the order of ascending and descending series. The size of the comparison card was also systematically varied for each series.

RESULTS

The mean judgments as well as the average constant error (CE) and average IU for each age level are presented in Table 1. A separate analysis

TABLE 1
MEAN PSE, CE, AND IU FOR EACH AGE LEVEL

Distance of Standard	Age	PSE Size of Standard			Over-all Average for All Standards	
		5 cm.	6 cm.	7 cm.	CE	IU
2 m.	17	5.40	6.44	7.16	.33	.72
2 m.	12	5.82	6.81	7.51	.71	.84
2 m.	7	5.71	6.93	7.67	.77	.95
2 m.	5	6.10	6.94	7.94	.99	1.42
6 m.	17	5.29	6.36	7.22	.29	.48
6 m.	12	5.39	6.52	7.27	.39	.60
6 m.	7	5.34	6.52	7.25	.37	.86
6 m.	5	5.44	6.42	7.15	.33	1.16

CHILD DEVELOPMENT

of variance was computed for each size standard. Although a significant F ($p < .01$) was obtained for distances (2 vs. 6 m.) in each of the three analyses, the F for age levels was significant ($p < .05$) only for the 5 cm. standard. The judgments of the standards at 6 m. were more accurate than those at 2 m. In addition, there were no consistent differences between the various age levels at the 6 m. distance.

In order to evaluate the over-all performance for all three standards at the 2 m. distance, a separate analysis was performed. The CE, which is obtained by subtracting the PSE from the size of the standard, was used in this analysis (8). It should be noted that differences in average CE are identical in numerical value to differences in average PSE for each age level. However, the use of the CE makes comparison of accuracy with the different size standards possible. This analysis appears in Table 2.

TABLE 2
ANALYSIS OF VARIANCE OF CE WITH THE STANDARDS AT 2 M.

Source	df	M.S.	F	p
Size (S)	2	.716	1.81	$>.10$
Age (A)	3	2.140	5.42	.005
S $\times$ A	6	.088	.22	
Individuals	114	.394		
Total	125			

A significant F ($p < .005$) was obtained for age levels. The magnitude and statistical significance of the differences between the mean CE for the four age levels appear in Table 3. The 17-year-olds were more accurate than all the other age groups ($p < .01$) when the standard was at 2 m. None of the younger age groups differed significantly from one another. There was also no significant difference in the accuracy of judgment of the three standard objects.

TABLE 3
DIFFERENCES IN OVER-ALL CE AMONG THE FOUR AGE LEVELS

Age	STANDARD AT 2 M.			STANDARD AT 6 M.		
	12	7	5	12	7	5
17	.38*	.44*	.66*	.10	.08	.04
12		.06	.28		-.02	-.06
7			.22			-.04

* $p < .01$.

In the analysis of the IU, nonparametric procedures were utilized because the distributions were skewed. By means of the sign test (6), a significant difference ($p < .05$) was obtained between the mean IU at 2 m. and that at 6 m. Thus, there was a reduction of equivalence at 6 m. An analysis of the differences between age levels, using White's extension of the method of Wilcoxon (6), indicated that both the 17- and 12-year-olds had a significantly lower IU than the 5-year-olds at 2 m. ($p < .05$) and also at 6 m. ($p < .01$). In addition, the 17-year-olds differed significantly ($p < .01$) from the 7-year-olds at 6 m.

Although no differences were found between the various age groups in accuracy, i.e., in size constancy, at the 6 m. distance, the groups differed considerably in IU at this distance. The implications of this finding as well as the other results will be discussed in the following section.

DISCUSSION

The above results clearly show that young adults have better size constancy than children when the distance between standard and comparison objects is relatively large and that these differences disappear when the distance between the objects is reduced. In the light of these findings, it is possible to reinterpret some of the contradictory data in the size constancy literature. Above all, it is now clear that the conception of size constancy as a unitary phenomenon is unwarranted.

Upon re-examination of the evidence on both sides of the so-called nativism-empiricism controversy, it becomes apparent that supporters of the empiricist position (1, 2, 9) have tended to emphasize those results within an experiment which demonstrated superiority of the adults over the children. At the same time, they tended to neglect any indications of similarity between the age groups. The nativists did exactly the reverse (3, 4).

One of the factors which seems to account for variations in the results of previous investigations is that of the distance between the two stimulus objects. For example, Beryl's data (1), which have often been used to support a learning approach, clearly indicate that the differences in size constancy between age levels decrease as the distances between test objects decrease. When this distance is only 2 m., there are no differences in judgment after 3 years of age. Zeigler and Leibowitz (9) also tend to stress the differences between adults and children when the distances between test objects are large, i.e., 25 ft. and over, but ignore the equal size constancy of the children and adults when the distance between the test objects is small, i.e., 5 ft.

The nativists often cite the investigations by Burzlaff (3) in which near perfect size constancy was found at all age levels. This performance has often been attributed to the method he utilized, which consisted of the

DEVELOPMENTAL, PSYCHOSOCIAL, AND EDUCATIONAL FACTORS IN CHILDREN WITH NONORGANIC ARTICULATION PROBLEMS¹

RUTH FITZSIMONS

*Warwick School Department
Warwick, Rhode Island*

Nonorganic, articulatory problems, involving sound substitutions, omissions, distortions and additions are the most common of all speech disorders. One research study (10) has placed the incidence of articulatory problems at 50 per cent of all speech difficulties which are represented among school children. Investigation in this area is justified by the numerical supremacy of the problem (13) and by the comparative paucity of available research (9, 13).

An articulation problem is a manifestation that the child's speech development is impaired and that his language progression is protracted. The primary problem investigated in this study may be stated as follows: Do children who have nonorganic, articulatory problems differ from children who speak normally in ways other than in the speech disorder syndrome? The rationale of the present investigation was that complete understanding of the child and his articulatory problem comes only by means of an organismic study of the child in the light of his developmental, educational, psychosocial, and speech attributes.

SUBJECTS

The children for this study were drawn from a sample population of 1,477 first grade children in the public school system of Warwick, Rhode Island. The subjects were divided into two groups consisting of 70 children each. The control group was comprised of 70 children whose speech was identified as normal, the experimental group, of 70 children whose speech

¹ This article is based on an Ed.D. dissertation completed under the direction of Dr. Wilbert L. Pronovost, Director of the Boston University Speech and Hearing Center.

CHILD DEVELOPMENT

handicaps were diagnosed as diffuse, nonorganic, articulatory problems. The group members were screened and found to be free from physical anomalies of the speech and otic mechanisms. On the basis of the results obtained from sweep frequency tests of hearing acuity, only those children whose hearing acuity was found to be within the normal limits of 15 decibels were included in the normal and articulation samples.

The matching criteria for the two groups were CA within four months; IQ, within five points; sex and school locale. The articulation group included 40 males and 30 females, ranging in age from 6 years, 2 months, to 7 years, 10 months (median, 6 years, 7 months). The normal group consisted of 40 males and 30 females who ranged in age from 6 years, 2 months to 7 years, 9 months (median, 6 years, 8 months). For the articulation group, IQ's ranged from 83 to 123 (median, 108), and from 81 to 120 for the normal group (median, 107).

PROCEDURE

Historical case study data were obtained on the 140 subjects by means of a rapport-type, personal interview of the mother by the experimenter at the home of each child. The following data were secured:

1. Child's date of birth and chronological age
2. Type of pregnancy
3. Type of birth
4. Type of labor
5. Child's age at the time of weaning from the bottle to the cup
6. Child's age at the time toilet training was implemented
7. Child's age at the time toilet training was accomplished
8. Age of onset of speech:
 - a. word
 - b. phrase
9. Age of onset of locomotion (walking)
10. Illnesses which occurred during the period of neonate through the sixth year of life
11. Presence or absence of the following childhood problems as chief complaints at the time of the interview:
 - a. destructiveness
 - b. eating and food problems
 - c. enuresis
 - d. fears
 - e. finger nail biting
 - f. the hurting of pets
 - g. jealousy
 - h. lying
 - i. nervousness
 - j. refusal to obey
 - k. showing off
 - l. shyness
 - m. sleeplessness
 - n. temper tantrums
 - c. thumb sucking

The reading readiness status of each subject was determined by means of the Metropolitan Reading Readiness Test, Form R (5), administered in September, 1953. In February, 1954, each child in the first grade sample population was administered the Kuhlmann-Anderson Group Test of Intelligence (6). At the close of the academic year in May, 1954, the achievement level of each child was determined by the Metropolitan Achievement Tests (4). Final report card grades were also available for analysis. The Vineland Social Maturity Scale (3) was used as a part of the parental interview.

The Children's Apperception Test was administered and the protocols interpreted on a descriptive level, noting the average number of words per protocol and the absence or presence in each protocol of the following dynamisms: orality, hostility directed toward parental figures, perception of parental figures as authoritarian, aggressive phantasies, sibling rivalry, fears and anxieties, significant conflict between autonomy-compliance, and positive or negative outcome. Three judges, one the experimenter, analyzed the protocols of 10 per cent of the subjects, selected by ordering the children in the total sample from highest to lowest IQ and then choosing every tenth child. No significant difference was found among judges in the problem areas identified.

Each child was rated in May, 1953, by his teacher on A Teacher Administered Rating Chart for Talks at the Primary Level (11) with respect to knowledge of topic, organization and content of topic, appearance, voice, and amount of participation. This study was concerned with the total score obtained by each child.

TABLE I
COMPARISON OF ARTICULATION AND NORMAL SAMPLES ON
DEVELOPMENTAL FACTORS

<i>Developmental Factors</i>	<i>Articulation Sample (N=70)</i>	<i>Normal Sample (N=70)</i>	<i>Chi Square</i>
Abnormal birth	16	7	4.21*
Bottle feeding	68	62	3.87*
Weaning before 6 months of age	28	6	18.80*
Toilet training before 9 months of age	37	24	4.91*
Accomplishment of toilet training before 2 years of age	22	11	4.80*
Occurrence of childhood diseases before or at the age of 3 years	40	20	11.67*
Walking before 12 months of age	18	33	6.94*
Talking (words) before 12 months of age	17	36	10.96*
Talking (phrases) before 18 months of age	17	36	10.96*

* Significant at .05 level.

CHILD DEVELOPMENT

Null hypotheses pertaining to the discrete data of this study were tested by chi square for correlated samples. The continuous data were subjected to the *t* test for related measures.

FINDINGS

Table 1 reveals that significantly more members of the articulation group experienced abnormal birth conditions, bottle feeding, the imposition of early weaning (prior to 6 months of age), the early occurrence of childhood diseases, early age of implementation and accomplishment of toilet training, and delay in both locomotive and communicative attempts.

TABLE 2
COMPARISON OF ARTICULATION AND NORMAL SAMPLES ON
PRESENCE OF CONDUCT AND HABIT DISORDERS

<i>Conduct and Habit Disorders</i>	<i>Articulation Sample (N=70)</i>	<i>Normal Sample (N=70)</i>	<i>Chi Square</i>
Destructiveness	18	2	14.93*
Eating and food problems	39	2	47.22*
Enuresis	9	3	3.28
Fears	21	10	5.01*
Finger-nail biting	7	11	1.10
Hurting pets	1	0	.11
Jealousy	6	3	5.08*
Lying	6	1	3.77
Nervousness	53	22	27.60*
Refusal to obey	27	7	15.54*
Showing off	30	6	21.54*
Shyness	31	13	10.74*
Sleeplessness	6	1	3.77
Temper tantrums	35	8	24.47*
Thumb sucking	34	15	11.33*

* Significant at .05 level.

Table 2 indicates that destructiveness, eating and food problems, fears, jealousy, nervousness, refusal to obey, showing off, shyness, temper tantrums, and thumb sucking existed to a significantly greater extent in the children who presented articulation disorders. No significant differences were apparent between the two groups in these psychologic problems: enuresis, finger-nail biting, hurting pets, lying, and sleeplessness.

The incidence of unsatisfactory report card grades in health habits, work habits, and language as well as unsatisfactory scores in reading readiness status and reading grade placement was significantly greater within

RUTH FITZSIMONS

TABLE 3

COMPARISON OF ARTICULATION AND NORMAL SAMPLES ON
EDUCATIONAL FACTORS

<i>Educational Factors</i>	<i>Articulation Sample (N=70)</i>	<i>Normal Sample (N=70)</i>	<i>Chi Square</i>
Below average reading readiness status	25	12	6.21*
Unsatisfactory reading grade placement on achievement test	26	12	7.08*
Nonpromotion to grade 2	16	9	2.39
Unsatisfactory report card grade in handwriting ..	12	7	1.52
Unsatisfactory report card grade in health habits ..	16	3	10.29*
Unsatisfactory report card grade in language	17	1	14.34*
Unsatisfactory report card grade in reading	17	8	3.94*
Unsatisfactory report card grade in social habits ..	15	7	3.45
Unsatisfactory report card grade in work habits ..	19	6	8.22*

* Significant at .05 level.

the speech handicapped sample. No relationship was found between membership in the articulation group and frequency of nonpromotion to grade two, unsatisfactory grades in handwriting or in social habits. (See Table 3.)

Table 4 summarizes the results of the analysis of the CAT protocols. The articulation group exceeded to a significant degree the normal group in aggression, fears and anxieties, and perception of parents as authoritarian. The mean number of positive outcomes was significantly greater for the normal children.

TABLE 4

DIFFERENCES BETWEEN THE ARTICULATION AND NORMAL GROUPS ON
THE CAT AND ON ORAL PARTICIPATION

	$M_a - M_n$	t	p
<i>CAT Dynamisms</i>			
Aggression	1.30	3.74	.01
Conflict between autonomy-compliance	.50	1.81	ns
Fears and anxieties	.30	8.10	.01
Hostility directed toward parental figures ...	.29	2.00	ns
Mean no. of words per protocol	7.26	1.86	ns
Orality	.10	.04	ns
Parents perceived as authoritarian	1.20	4.26	.01
Positive outcomes	-1.26	3.04	.01
Sibling rivalry	.03	.62	ns
<i>Total Oral Participation Score</i>	-6.04	4.17	.01

CHILD DEVELOPMENT

The mean total oral participation score was significantly higher for the normal group (Table 4). There were no significant differences on the Vineland Social Maturity Scale.

DISCUSSION

The thesis that language problems appear to be psychogenic in etiology has been given empirical support (7). Mowrer (8), considering speech development in the light of psychoanalytic and modern learning theories, emphasized the importance of opportunities for the child to attend to auditory stimuli under conditions which are conducive to pleasurable associations. The child's interactions with the important people in his psychosocial milieu are the origins out of which language learnings develop. Too early and too rigid an imposition of demands for weaning and toilet training are considered to portend subsequent anxiety and personality problems (2, 12). An audit of the anamneses of the articulation sample reveals that this group experienced an early initiation of both weaning and toilet training in addition to the early onset of childhood illnesses. Such a combination of factors may be considered to constitute negative valences in the life experiences of the children who comprised the articulation sample during the decisive period of their language learning. The articulation group tended to be delayed in both their communicative and locomotive attempts, talking and walking at later ages than did the normal sample. The articulation group members were reported to be bottle fed; because the difference between the articulation sample and the normal sample was not great, no conclusions concerning this variable can be drawn. Historical data also revealed that the incidence of abnormal birth conditions was significantly greater in the articulation sample.

The prevalence of tension-reducing symptoms in the articulation subjects is indicated by the significantly greater number of such behavior and conduct disorders as destructiveness, eating and food problems, fears, jealousy, nervousness, refusal to obey, showing off, shyness, temper tantrums, and thumb sucking.

With regard to the inferior status of the children in the articulation group in language, in health habits (which included, among other items, the avoidance of putting things into the mouth), and in oral participation, it may be assumed that the teachers, in their evaluations, were influenced negatively by the presence of articulatory errors in speech. It is interesting to note that each of the factors has an oral component.

The CAT protocols with their references to destruction, injury, and mutilation, their projection of fears and anxieties, their perceptions of parents in authoritarian, demanding roles, and their negative outcomes all suggest a basically negative orientation in these subjects.

The failure to find differences between the normal and articulation subjects in social competence may be explained, in part, by the fact that

the child with the articulation problem finds in the acquisition of social skills and proficiencies, particularly as reinforced by parents, a modality for gaining recognition.

The children who presented diffuse, nonorganic articulation problems differed from the children who spoke normally in ways other than in the speech disorder syndrome. The findings of this study suggest that a relationship may exist between a nonorganic articulation disorder and psychological factors. Two tentative hypotheses may be postulated. A nonorganic speech problem may be satisfying a need for the reduction of internal tension. Such a speech disorder may arise when the mechanism of identification (vital to language learnings) fails to prosper within the child whose early psychosocial milieu is surcharged negatively. Additional research on the relationship between a diffuse nonorganic articulatory disorder and psychological aspects is indicated.

The findings and implications of this study do serve to support Travis' (14) promulgation relative to the limitations of an essentially symptomatic treatment approach in speech therapy. In recognition of the possible relationship between a nonorganic articulation problem and psychologic factors, it appears that an efficacious therapeutic process requires a complete and continuous study of the child, his total needs, his adaptive mechanisms, and his interpersonal relationships as well as a consideration of his articulatory aberrations.

SUMMARY

The purpose of this study was to determine if speech handicapped children differed from normal speaking children in ways other than in the speech disorder syndrome. Data were gathered from a study of two matched groups: a control group of 70 children with normal speech and an experimental group of 70 children with diffuse, nonorganic articulatory problems. Historical case study data were secured. Projective and nonprojective test results were analyzed for both groups. The articulation sample revealed more deviation in the developmental, psychosocial, and educational variables tested than did the normal speaking sample. The major findings were:

1. Members of the articulation group experienced significantly more abnormal birth conditions, bottle feeding, early weaning, the occurrence of childhood diseases during the first three years of life (a consequential period for language learning), the implementation and the accomplishment of toilet training at an early age, as well as delay in both locomotive and communicative attempts.
2. The conduct and habit disorders of destructiveness, eating and food problems, fears, jealousy, nervousness, disobedience, showing off, shyness, temper tantrums, and thumb sucking existed to a significantly greater extent in the speech handicapped sample.

CHILD DEVELOPMENT

3. Unsatisfactory report card grades in health habits, work habits, and language as well as unsatisfactory scores on reading readiness status and reading grade placement were found more frequently within the speech handicapped sample.

4. Aggression, fears and anxieties, and perception of parental figures as authoritarian appeared significantly more often in the CAT protocols of the articulation sample.

5. The mean number of positive CAT protocol outcomes for the normal sample was significantly greater than that for the articulation group.

6. The average total oral participation scores of the normal group were significantly greater than those of the speech handicapped group.

The present investigation suggests the existence of a relationship between a diffuse, nonorganic articulatory disorder and psychosocial factors, indicating a need for additional research. These findings preclude the efficacy of an essentially symptomatic treatment approach in speech therapy with children who present diffuse, nonorganic articulatory problems.

REFERENCES

1. BELLAK, L., & BELLAK, SONYA S. *Children's Apperception Test*. New York: C.P.S., 1946.
2. BLAU, A., & HULSE, W. C. Preventive psychiatry in the care of the newborn and infant. Paper read at Amer. Psychiat. Ass., Los Angeles, May, 1953.
3. DOLL, E. A. *Vineland Social Maturity Scale*. Philadelphia: Educational Test Bureau, 1948.
4. HILDRETH, GERTRUDE H. *Metropolitan Achievement Tests*. New York: World Book, 1948.
5. HILDRETH, GERTRUDE H. *Metropolitan Readiness Tests*. New York: World Book, 1949.
6. KUHLMANN, F., & ANDERSON, ROSE G. *Kuhlmann-Anderson Test*. (6th Ed.) New Jersey: Personnel Press, Inc., 1952.
7. MCCARTHY, DOROTHEA. Language disorders and parent-child relationships. *J. Speech Hearing Disorders*, 1954, 19, 514-523.
8. MOWRER, O. H. Speech development in the young child: 1. The autism theory of speech development and some clinical applications. *J. Speech Hearing Disorders*, 1952, 17, 263-268.
9. POWERS, MARGARET H. Functional disorders of articulation—symptomatology and etiology. In L. E. Travis (Ed.), *Handbook of speech pathology*. New York: Appleton-Century-Crofts, 1957. Pp. 707-768.

RUTH FITZSIMONS

10. PRONOVOST, W. L. A survey of services for the speech and hearing handicapped in New England. *J. Speech Hearing Disorders*, 1951, 16, 148-156.
11. READ, ELIZABETH C. A teacher administered rating chart for talks at the primary level. Unpublished Master's thesis, Boston Univer., 1953.
12. SILVERBERG, W. V. *Childhood experiences and personal destiny*. New York: Springer, 1952.
13. SPRIESTERBACH, D. C. Research in articulation disorders and personality. *J. Speech Hearing Disorders*, 1956, 21, 329-335.
14. TRAVIS, L. E. A point of view on speech correction. *Quart. J. Speech*, 1936, 22, 57-61.

SELF-CONCEPTIONS, MOTIVATIONS, AND INTER- PERSONAL ATTITUDES OF EARLY- AND LATE-MATURING GIRLS¹

MARY COVER JONES *and* PAUL HENRY MUSSEN

University of California, Berkeley

"The changing body and the changing self" is a phrase associated with adolescent development (18). It suggests that the shaping into mature form of the childhood body pattern is accompanied by new self-concepts. These altered attitudes toward the self reflect at least in part the youth's response to his physical metamorphosis.

What "growing-up" connotes for the individual adolescent depends upon a complex of psychobiological factors. One of the most important of these is rate of physical maturation. Adolescent growth may be relatively regular and even, or it may be uneven or abrupt. The timing of puberty, in relation to social norms of the peer group, may present problems of special importance for some adolescents.

Previous reports of systematic comparisons between the behavior and personality characteristics of early- and late-maturing adolescents have indicated that acceleration in growth tends to carry social advantages for boys (9, 10) but disadvantages for girls (7). At their peak of growth, early maturing girls are not only taller than their girl classmates but are actually taller than most of the boys in their class (14). They are conspicuously large at a time when physical size is not an asset for girls in our culture. Many girls consider tallness to be a physical stigma (15). At the end of adoles-

¹ The TAT data for this study were obtained by Harold E. Jones in connection with a test program at the Institute of Child Welfare.

CHILD DEVELOPMENT

cence the early-maturing are no longer taller than their age-mates, but in body proportion they tend to have a broad and stocky build (13), less attractive (in terms of current feminine standards) than the more slender physique of the late-maturing.

Among boys, ascendance in size and musculature is an asset because of our cultural values and the functional advantages of such a build for athletic prowess. This more favorable status is indicated in observational records for early-maturing boys. Staff members rated them as physically more attractive and better-groomed than the late-maturing, and in social situations they were more poised and matter-of-fact, and less attention-seeking (9).

In contrast, both classmates (16) and adult observers (12) saw the early-maturing girls as relatively submissive, listless or indifferent in social situations, and lacking in poise (7). Such girls have little influence upon the group and seldom attain a high degree of popularity, prestige or leadership (8).

The girls in the slower-maturing classification were seen as relatively more outgoing and more assured. They were eager, animated, peppy, and talkative. This behavior seems to be acceptable among girls since those who exhibit it are also described as confident and having leadership abilities (7). While the same characteristics of expressiveness are attributed to slow-growing boys, it is associated in their case more specifically (and especially in later adolescence) with show-off behavior, affectation, and tenseness (6, 9).

In accounting for these sex differences in the response to early or late puberty, we may note that although early-maturing boys have physical advantages over other boys and are socially in step with girls, the girl who develops earlier than her classmates may be temporarily isolated. H. E. Jones has expressed this as follows:

The early-maturing girl quite naturally has interests in boys and in social usages and activities more mature than those of her chronological age group. But the males of her own age are unreceptive, for while she is physiologically a year or two out of step with the girls in her class, she is three or four years out of step with the boys—a vast and terrifying degree of developmental distance (7, p. 78).

A study of responses to the Thematic Apperception Test, given to members of the Adolescent Growth Study when they were seniors in high school, yielded a somewhat unfavorable psychological picture for the late-maturing boys. Compared with their early-maturing peers, they showed greater evidence of negative self-concepts, prolonged dependency needs, feelings of rejection by others, rebellious attitudes toward parents, and strong affiliative needs. These findings were in agreement with evidence from other sources (10).

A similar TAT comparison of early- and late-maturing girls should be expected to show results different from those obtained for boys. Thus, it

might be expected that early-maturing girls would reveal negative self-feeling and less satisfactory interpersonal attitudes.

PROCEDURE

The present study, paralleling that for boys, was designed to investigate the relationship between maturational status and self-conceptions, motivations, and interpersonal attitudes in a normal public school sample of girls (5). Personality assessment was made on the basis of their responses to the Thematic Apperception Test (TAT).

The 34 17-year-old girls of this investigation constitute approximately the 20 per cent at each extreme of the total sample of the Adolescent Growth Study, selected on the basis of their physical maturity status as determined by X-rays of the wrists and hands (1). Sixteen had been among the most consistently accelerated over a four-year period during adolescence; the other 18 were among the most consistently retarded. All of the subjects took the TAT at around age 17 when they were seniors in high school.

The TAT consisted of 18 pictures: nine from the Murray set which is now standard (cards 1, 5, 6, 7BM, 10, 11, 14, 15, 17); five pictures from the set generally used in 1938 when these data were collected (a man and woman seated on a park bench; a bearded old man writing in an open book; a thin, sullen, young man standing behind a well-dressed older man; a tea table and two chairs; an abstract drawing of two bearded men); and four cards not in the Murray series (a madonna and child, the nave of a large church, a dramatic view of mountains, a boy gazing at a cross which is wreathed in clouds).

The tests were administered individually. Each card was projected on a screen while the subject told a story which was recorded verbatim. Standard instructions were given for the Murray cards, and subjects were asked to describe the feelings elicited by the other four pictures. Most of the stories were brief.

The scoring scheme involved counting the relevant needs, press, and descriptions of the heroes of the stories, the assumption being that the storyteller has identified with the hero; the hero's needs are the same as the girl's; the press that impinge upon the hero are the ones that affect the girl telling the story. A total of 20 needs, press, and descriptive categories, each defined as specifically as possible, was developed in the analysis of the protocols. A score for each subject for each TAT category was derived by counting the number of stories in which it appeared. Table 1 presents a list of the categories used, together with definitions of these categories.

To test the reliability of this analysis, one of the authors (PM) and another psychologist² independently scored 15 complete protocols (300

² We are indebted to Dr. Virginia B. Ware for her participation in this aspect of the study.

stories). The percentage of interrater agreement was 90, computed by the usual formula (number of agreements divided by number of agreements plus number of disagreements).

In order to eliminate bias, the scoring used in the present study was done "blind," that is, independently of knowledge of the subject's maturational status.

RESULTS

Frequency distributions of the scores of all subjects were made for all the TAT variables. Each distribution was then dichotomized at the point which most nearly enabled the placing of half of the 34 subjects above, and half of them below, the dividing point. Subjects having scores above this point were considered high in this particular variable; those with scores below this point were considered low in this variable.

Table 1 lists the TAT variables together with the number of late- and early-maturers with high scores in each variable. The exact probabilities of obtaining these distributions of high and low scores in the two groups (or all other possible more extreme sets), calculated in accordance with Fisher's method (3), are given in the fourth column. The last column gives the levels of significance of the differences between early- and late-maturing boys on these same variables, previously reported by the authors (10).

As may be seen from Table 1, early- and late-maturing boys differed from each other on many more characteristics than the two groups of girls did. The boys' groups were significantly different from each other, at the 5 per cent level or better, in six of the 20 variables scored, while the early- and late-maturing girls differ significantly in only two of the variables (*negative characteristics* and *n Recognition*). It should be noted, however, that the *direction* of the differences tended to be the same, rather than reversed, in the two sets of data. For example, the following similarities may be noted:

1. In this list of characteristics a significantly greater proportion of late-maturing girls than of early-maturing girls have high scores on *negative characteristics*. This finding is similar to that found in the comparison of early- and late-maturing boys. For girls, it is contrary to expectation.

2. The differences between early- and late-maturing girls in respect to *p Dominance* and *p Rejection* are similar to those for early- and late-maturing boys in these variables. These may be interpreted to indicate slightly poorer parent-child relationships among the late-maturing.

3. The early- and late-maturing boys differ significantly on *n Autonomy 1*, suggesting a greater tendency for the late-maturing to avoid or defy authority. The differences are in the same direction for girls, but are not significant.

4. Similar results for boys and girls in *n Succorance* may be interpreted as showing some tendency for stronger dependency needs in the late-maturing.

TABLE I

NUMBER OF EARLY- AND LATE-MATURERS SCORING HIGH
IN TAT VARIABLES

<i>TAT Variable (with Definition)</i>	<i>High Early- Maturers</i>	<i>High Late- Maturers</i>	<i>p</i>	<i>p</i> (boys)
<i>Negative Characteristics</i> —H is described in negative terms (e.g., imbecile, weakling, fanatic)	3	13	.002	.01
<i>p Dominance 1</i> —H forced by parents to do something he doesn't want to	2	7	.08	.09
<i>p Dominance 2</i> —H prevented by parents from doing something he wants to	3	6	—	—
<i>p Dominance 3</i> —Total instances of H's being forced by parents to do something and/or prevented from doing something	4	10	.07	.11
<i>p Rejection</i> —H rejected, scorned, or disapproved by parents or authorities	3	7	.21	.03
<i>n Aggression 1</i> —H is aggressive in physical, asocial way	8	9	—	.02*
<i>n Aggression 2</i> —H is mad at someone, argues	12	14	—	.10*
<i>n Aggression 3</i> —Total of all H's aggressive actions	5	8	—	.10*
<i>n Autonomy 1</i> —H leaves home	8	10	—	.20
<i>n Autonomy 2</i> —H disobeys or defies parents	4	5	—	.11
<i>n Autonomy 3</i> —Total of instances in which hero leaves and/or defies his parents	5	7	—	.02
<i>n Affiliation 1</i> —H establishes good relations with his parents	9	8	—	—
<i>n Affiliation 2</i> —H falls in love, has a romance, marries	11	7	.08	.05*
<i>n Affiliation 3</i> —Total instances in which H establishes and/or maintains friendly relations	7	6	—	.11*
<i>n Succorance</i> —H feels helpless, seeks aid or sympathy	5	10	.16	.06
<i>p Nurturance 1</i> —H is helped, encouraged, or given something by parents	5	7	—	.18
<i>p Nurturance 2</i> —H is helped, encouraged, or given something by someone else (not parents)	5	10	.16	.02
<i>n Achievement</i> —H attempts to attain a high goal or do something creditable	7	13	.09	—
<i>n Recognition</i> —H seeks fame and/or high prestige status	3	12	.01	—
<i>Denial of Feeling</i> —S states that picture elicits no thoughts or feelings	10	9	—	.06

* Differences are in the opposite direction for boys as compared with girls.

CHILD DEVELOPMENT

5. Similar results for boys and girls in *p Nurturance* (significant in one variable for boys) may also be interpreted as indirect indications of stronger dependency needs in the late-maturing.

The chief differences between the sexes are as follows:

1. With respect to *n Aggression* more early- than late-maturing girls show "argumentative aggression," but the two groups of girls do not differ in physical aggression. On the other hand, more early-maturing than late-maturing boys show high degrees of both kinds of aggression.

2. On one category of *n Affiliation* (involving romantic love) higher proportions of high scores are shown for early-maturing girls as contrasted with their late-maturing peers. The differences between the early-maturing and the late-maturing boys are in the opposite direction for this category.

3. The variables *n Achievement* and *n Recognition* do not differentiate the two groups of boys. Among girls scores are higher for the late-maturing, very significantly so in the case of *n Recognition*.

4. *Denial of feeling* does not differentiate early- and late-maturing girls but tends to yield higher scores for early-maturing boys.

DISCUSSION

The failure of the TAT data to support observational findings, especially with reference to the variable, negative characteristics, might be accounted for in a number of ways. Some writers report that in many cases thematic fantasies and manifest behavior operate independently and are even negatively related (11). If we assume this to be the case for our subjects, no further explanation would be needed. But there is also evidence from the literature that, for some groups, TAT findings and overt behavior may be congruent (4). Our data on boys are in line with this assumption, since, according to observational ratings, late-maturing boys tend to be socially disadvantaged, and, according to the TAT, personally more maladjusted.

The findings for girls are quite different, however. The early-maturing received more unfavorable ratings from both peers and adult observers on many characteristics. But in the TAT they appear to be somewhat better adjusted than their late-maturing peers. This discrepancy between observers' ratings and the picture derived from the personality tests may stem partly from the fact that the reported observational records represented an average of repeated ratings taken over a period of time (from 11 to 17 years) while the TAT stories were collected at the end of this period.

Girls who enter puberty early would be expected to have more difficulties in personal-social relations when they are out of phase with their group. However, after the peer group "catches up," these difficulties would be reduced. By the end of senior high school maturational discrepancies, and social distance due to this factor, would be less marked. It is also possible that even a slight improvement in status would bolster morale and be

reflected in a projective technique designed to register attitudes and self-concepts.

There is some slight evidence of a trend toward improved social status for the early-maturing in observational ratings over the seven-year period. Twenty-five items concerned with appearance, emotional tone, social participation, responsiveness, and assurance were used in the comparison. Three of these reflected an improved status at the twelfth grade level for early-maturing girls (2). In two of these, "laughing" vs. "sober" and "sociable" vs. "unsociable," the accelerated girls, while still rated lower than the late-maturing, had improved sufficiently so that the differences between the two groups were no longer significant. But for one important characteristic, "popular" vs. "unpopular," the average ratings for the accelerated girls were now actually slightly higher than for the late-maturing, though the differences were not significant. This last year at high school was the only period when the early-maturing girls were rated by observers as above average in popularity.

It is conceivable that other improvements in social relationships were undetected because of the "halo effect" which, in spite of precautions, may have influenced observers who had rated these same adolescents in earlier years. It is not unlikely that if these girls of more mature status had been observed in social groups of their own choosing (presumably outside of school) the behavior picture might have been more favorable.

It may be noted that over the seven-year period the observational records received little corroboration from a self-report inventory (17). Although differences were not consistent in all categories, the early-maturing girls tended to score more favorably than the slow-maturing on "total adjustment," and also on family adjustment and feelings of personal adequacy. These data from the self-report inventory seem to be generally consistent with the findings from the TAT.

However, we may note that in both the inventory and the TAT the early-maturing girls appear in a somewhat better light than in their reputation scores or in ratings by adult observers. In some individual cases a favorable self-report score should not be taken at face value, in view of the tendency for some individuals to cover up or deny their deficiencies.

The only other variable which yields a significant difference between the maturity groups is the category *n* *Recognition*, late-maturing girls manifesting a greater desire for personal recognition. The results for *n* *Achievement*, though not showing significant differences, tend to support these findings. Other data for this group of girls would lead us to expect this relationship between maturity status and desire for recognition. Late-maturing girls were rated by adult observers as attaining higher prestige, showing more leadership, and having greater stimulus value than their early-maturing peers (7). They were also mentioned more frequently in the high school daily paper over a three-year period and were elected to more offices in extra-curricular activities (8). The late-maturing girls' lead-

ership abilities, their greater social participation, and their apparent social success may have been more closely related to desires for recognition and achievement in the social sphere than to a need for affiliation (8).

It should be noted that, among boys, *n Achievement* and *n Recognition* were not significantly associated with rate of physical maturation. Perhaps this is due to the fact that for boys in our culture the pressures to strive for achievement and personal recognition are powerful and pervasive; hence, the boy's physical status may have little influence on his acquisition of strong achievement and recognition needs. Since these cultural pressures are undoubtedly less severe for girls, the strength of these personal needs may be more influenced by such factors as rate of physical maturation.

As we have pointed out in an earlier article (10), the relationship between physical status and psychological characteristics in boys is by no means simple. The evidence of the present study indicates that this relationship is even more complex in the case of girls. While the TAT analysis reported in this study suggests that early-maturing girls have fewer negative self-concepts and fewer needs for personal recognition, the results must be interpreted very cautiously. Since only two variables were found to be significantly related to physical status, it is obvious that many psychological and social factors are more important than rate of maturing in determining girls' self-concepts and personality characteristics. Furthermore, these data, considered together with the data from earlier studies on girls (7, 8), suggest that the rate of maturation may affect overt behavior and covert characteristics in different—sometimes seemingly contradictory—ways.

It is also possible that, at least for girls, early- or late-maturing means different things at different stages of adolescent development. It has been proposed that since girls who enter puberty early are out of step physically with both the boys and girls in their classrooms, they tend to be socially handicapped during early adolescence. We have assumed that this would carry emotional hazards, and evidence is available from observational data and reputation measures to indicate that this is the case (7).

However, the accelerated girl may gain assurance from knowing that she is on the way toward a goal which is a common task for all adolescents, that of being an adult. By the end of high school, many girls in this group were beginning to feel that they had made satisfactory progress toward this goal. If, in addition to this, she can cope with the problems of this period without too much stress, her self-esteem and feelings of adequacy may be enhanced. A resulting improvement in self-concepts may be reflected in the relative infrequency of negative characteristics in TAT stories.

In conclusion, it is evident that each individual's unique personality structure is determined by a complex of interacting variables, including rate of maturation. Comments made by these subjects as young adults indicate that they were aware of a variety of surface phenomena which affected their adolescent adjustment:

"High school is not a pleasant memory. I felt remote from my mother. If I could have talked to her, it would have helped" (a slow-maturer).

"I wasn't very happy in adolescence. My father was out of work. I felt inferior outside my own circle of friends—I always aimed to please" (a very popular late-maturing girl).

"I was slightly rattle-brained" (a popular late-maturing girl).

"I didn't have much fun in high school. I look forward to more happiness now than I did when I was in high school. I was an ugly duckling" (a slow-maturer who ascribed many negative characteristics to the hero).

"I seemed to be separated from friends in high school. I'm more outgoing now, less cautious and fearful" (accelerated girl).

"I was overweight and sensitive about it—now I take things more for granted" (accelerated girl).

"I had a feeling of being different when growing up" (accelerated).

"I felt stupid in school" (accelerated girl).

"I was very lacking in self-confidence in high school" (accelerated girl).

"I'm more optimistic now. I didn't know many people in high school. I would make an effort to get on with people if I had it to do over again" (accelerated).

Feelings of inadequacy and isolation are expressed by these girls and they are attributed to lack of mental ability, financial difficulties, separation from parents, poor social status, overweight, and unattractiveness. They are about equally common among those whose maturational status was at one extreme or the other.

It is obvious that the findings for this specific group of girls need to be particularized for each individual. These results might be modified also for girls in another geographical area or social level or in another generation. It is possible that school and community programs may be able to de-emphasize maturational status by providing an easier access to mixed social groups through classroom, extra-curricular, and recreational activities which cut across age classifications.

SUMMARY

The present study was designed to investigate the relationship between maturational status and TAT scores for a group of physically-accelerated as contrasted with a group of slow-developing girls from a normal classroom sample. The TAT protocols of 34 17-year-old girls—16 who had been consistently accelerated and 18 who had been consistently retarded—were analyzed according to a scoring scheme involving 20 needs, press, and descriptive categories.

The scores of early- and late-maturing in each of the categories were compared. Earlier reports (7, 8) had indicated that girls who reach puberty early are likely to be socially disadvantaged, at least until the rest of their age group "catch up" with them. It was assumed that this social disadvantage would be reflected in the TAT protocols and that differences

between the two maturity groups in self-concepts, attitudes, and motivations would be found. Analysis of the data of the present study found few striking differences between the two groups of girls. However, early-maturing girls had significantly lower scores on the category *negative characteristics*, indicating more favorable self-concepts. This finding is contrary to what might have been expected on the basis of observational ratings by adults and reputational ratings by classmates. On the other hand, the TAT results are in line with scores (total adjustment, self-adequacy, family adjustment) on a self-report inventory.

Late-maturing girls have significantly higher scores on *n Recognition*, which is corroborated by data from other sources.

When the differences between early- and late-maturing girls are compared with the differences between early- and late-maturing boys (10), they are found to be in the same direction more often than in the opposite. These findings are interpreted to indicate that late-maturing adolescents of both sexes are characterized by less adequate self-concepts, slightly poorer parent-child relationships, and some tendency for stronger dependency needs.

It has been emphasized that complex psychological and cultural factors as well as maturational status contribute to personality development and that the pattern of these influences varies for each individual.

REFERENCES

1. BAYLEY, N. Size and body build of adolescents in relation to rate of skeletal maturing. *Child Developm.*, 1943, 14, 51-89.
2. EVERETT, E. G. Behavioral characteristics of early- and late-maturing girls. Unpublished Master's thesis, Univer. of California, 1943.
3. FISHER, R. A. *Statistical methods for research workers*. (7th Ed.) Edinburgh: Oliver & Boyd, 1938.
4. HARRISON, R. The Thematic Apperception Test. In L. K. Frank *et al.*, Personality development in adolescent girls. *Mongr. Soc. Res. Child Developm.*, 1951, 16, 60-88.
5. JONES, H. E. Observational methods in the study of individual development. *J. consult. Psychol.*, 1940, 4, 234-238.
6. JONES, H. E. *Development in adolescence*. New York: Appleton-Century, 1943.
7. JONES, H. E. Adolescence in our society. In *The family in a democratic society: anniversary papers of The Community Service Society of New York*. New York: Columbia Univer. Press, 1949, Pp. 70-82.
8. JONES, M. C. A study of socialization at the high school level. In press.

9. JONES, M. C., & BAYLEY, N. Physical maturing among boys as related to behavior. *J. educ. Psychol.*, 1950, 41, 129-148.
10. MUSSEN, P. H., & JONES, M. C. Self-conceptions, motivations, and interpersonal attitudes of late- and early-maturing boys. *Child Develpm.*, 1957, 28, 243-256.
11. MUSSEN, P. H., & NAYLOR, H. K. The relationship between overt and fantasy aggression. *J. abnorm. soc. Psychol.*, 1954, 49, 235-240.
12. NEWMAN, F. B. The adolescent in social groups. *Appl. Psychol. Monogr.*, 1946, 9.
13. SHUTTLEWORTH, F. K. Sexual maturation and the physical growth of girls age 6 to 19. *Monogr. Soc. Res. Child Develpm.*, 1937, 2.
14. SHUTTLEWORTH, F. K. The physical and mental growth of girls and boys age 6 to 19 in relation to age at maximum growth. *Monogr. Soc. Res. Child Develpm.*, 1939, 4.
15. STOLZ, H. R., & STOLZ, L. M. Adolescent problems related to somatic variations. *Yearb. nat. Soc. Stud. Educ.*, 1944, 43, Part 1. Pp. 80-99.
16. TRYON, C. M. Evaluations of adolescent personality by adolescents. *Monogr. Soc. Res. Child Develpm.*, 1939, 4.
17. TRYON, C. M. *U. C. Adjustment Inventory I: Social and emotional adjustment*. Berkeley: Univer. of California Press, 1939.
18. ZACHRY, C. B. *Emotion and conduct in adolescence*. New York: Appleton-Century, 1940.

A REINTERPRETATION OF THE BARKER, DEMBO, AND LEWIN STUDY OF FRUSTRATION AND REGRESSION

JOHN M. DAVIS

Yale University School of Medicine

Barker, Dembo, and Lewin's (BDL) classic study of frustration and regression (1) has been widely discussed in psychological literature. Child and Waterhouse (CW) reinterpreted the experiment and have come to conclusions which differ from those of the original authors (2). In the present paper further analysis of the data will be presented in an attempt to evaluate the two interpretations.

BDL feel that psychological development can be characterized by structural changes in the life space such that there is increased differentiation of the person into subsystems and a change of organization of the person so that the subsystems are hierarchically related to one another rather than related by simple interdependence. BDL define regression as a primitivization of behavior: the subject's behavior becomes less organized and more like that of a less mature person. The BDL definition of regression differs therefore from the psychoanalytic definition in that the behavior elicited after frustration need not be specific responses that have been earlier manifested and abandoned by the organism. In the experiment 30 children (ages 3 to 5) were allowed to play with "average" desirable toys during a free play period. The next day they were allowed to play with some *very desirable* toys, but then both the toys of the first period and the children were moved to another part of the room and a wire screen partition was lowered so the children could see the very desirable toys, but could play *only* with those toys they had played with in the first situation. During both the free play and the frustration period the subject's behavior was recorded; later it was analyzed into a number of categories. The subject's play was rated on a constructiveness scale. The constructiveness scale measured the degree of organization, maturity, and differentiation of the play. It correlates highly with mental age and chronological age. The constructiveness of the play of children in the frustration situation was considerably lower than that in the free play situation.

In the analysis of the behavior of the children BDL differentiate between primary and secondary play. In primary play the child seems (to the observer) to be entirely absorbed in the play; in secondary play the child is playing but also doing something else at the same time. The constructiveness of secondary play is considerably lower than that of primary play.

Barker, Dembo, and Lewin interpret the experimental results in terms of their theoretical position outlined above. They feel that tension caused

by the "frustration" results in a dedifferentiation of the life space. According to Lewin's theory the degree of differentiation of a whole is inversely proportional to the strength of the pressure of the tension, when the tension is greater than limits set by the strength of the boundary of the natural parts of the whole. Lewin also feels that a decrease in the variety of behavior must occur when a part of the whole is kept in a fixed state because of the frustration.

Child and Waterhouse (CW) give the experiment a different interpretation (2). They point out that BDL's interpretation of the data in terms of dedifferentiation is circular, since no criterion for differentiation is present. They point out further that there is not, as BDL claim, a decrease in the variety of behavior during the frustration period; there is rather, in direct contradiction to BDL, an increase in the variety of behavior during the frustration period.

CW interpret the BDL study in terms of interference. That is, the lowered constructiveness of play in the frustration situation is explained by assuming that there is something about the frustrating situation which causes the subjects to make responses which interfere with the play responses. To illustrate the interference hypothesis let us take an example from another situation: Let us assume that there is a decrease in the quality of a man's driving when he is both driving the car and smoking a cigar, instead of driving the car only; that is, the responses of smoking the cigar interfere with those of driving the car.

CW presents the following evidence for the interference hypothesis: (a) During the frustration period, there was three times as much time devoted to nonplay activities as to play activities, and there was a greater proportion of secondary play. (b) With chronological and mental age partialled out, there is a correlation of -0.72 between the combined categories of escape and barrier behavior during the frustration period and constructiveness of play during this period. (c) With chronological and mental age partialled out, there is a correlation of -0.83 between constructiveness in the free play period and the combined categories of activity with the experimenter, island behavior, window behavior, and looking and wandering around during the free play period. (d) There is no significant difference in the constructiveness of secondary play between the play and the frustration period. This result would be predicted by the CW analysis because, in secondary play, there is already interference, so the extra interference in the frustration period would have no effect. The BDL interpretation cannot explain this fact.

However, it would be possible for BDL to counter the CW interpretation by some of the following counterarguments. BDL could argue in relation to CW points (a) and (b) that frustration "causes" the nonplay behavior and, therefore, the more frustrated subjects would engage in more nonplay behavior, i.e., more escape and barrier behavior. On point (c) BDL could argue that their theory makes no prediction in regard to indi-

vidual differences in the free play situation. Hence, the CW interpretation seems more reasonable intuitively, but their argument could be countered by the BDL group. What is needed to differentiate the two interpretations is more evidence.

PROBLEM

In their original paper, BDL give the score each subject made in each content category (1, Table 1, p. 85; Table 6, p. 119; Table 7, p. 123; Table 8, p. 133; p. 135; Table 9, p. 142; Table 20, p. 191; Table 21, pp. 191-192). It is possible, therefore, to go to these original data and analyze them in new ways.

The BDL interpretation and the CW interpretation would predict different correlations between these content categories. According to the BDL interpretation, those categories which are related to the strength of frustration would be negatively related to constructiveness. The BDL interpretation would therefore predict that the categories of mood satisfaction-unsatisfaction index¹ and escape behavior would correlate negatively with constructiveness. The CW interpretation would predict that those categories which imply the strongest interference states would correlate negatively with constructiveness, specifically, that the category of barrier behavior would correlate negatively with constructiveness. Barrier behavior would imply a strong interference state because it indicates that the subjects were concerned with the very desirable toys behind the barrier.

METHOD

Rank order correlations were calculated from the scores of each subject on each of five content categories. Each category was correlated with every other category. The matrix of correlations is listed in Table 1. The reliability of the various behavioral rating categories could not be calculated from the data available in the BDL report.

RESULTS

An inspection of the table of correlations shows that only those correlations predicted by the CW interpretation correlate significantly with constructiveness. Barrier behavior correlates $-.39$ with constructiveness and $-.50$ with maximum constructiveness. None of the correlations predicted by the BDL interpretations was significant: escape behavior correlated with constructiveness ($-.17$) and maximum constructiveness ($-.06$), and mood index correlated with constructiveness ($+.04$) and maximum con-

¹ BDL construct the mood satisfaction-unsatisfaction index from behavioral ratings of the following categories: singing and humming, friendly conversations, happy actions, unhappy actions, restless actions, knocking and kicking. The criteria for these categories and the construction of the index are discussed by BDL (1, pp. 187-204).

CHILD DEVELOPMENT

structiveness (+.29). In short, the evidence from the correlational analysis supports the CW interpretation and does not support the BDL interpretation.

TABLE I
INTERCORRELATIONS OF BEHAVIOR RATINGS

	C	MC	B	E	MI
C		.48*	-.39*	-.17	.29
MC	..		-.50*	-.06	.04
B	..	..		-.08	-.23
E	..	..	..		-.24

NOTE.—C, constructiveness; MC, maximum constructiveness; B, barrier behavior; E, escape behavior; MI, mood index.

* Significant at .05 level.

SUMMARY

Child and Waterhouse have reinterpreted the much discussed frustration and regression experiment of Barker, Dembo, and Lewin and have arrived at different conclusions from those of the original authors. Barker, Dembo, and Lewin rated the behavior of children in a free play situation and in a frustration situation on a number of content categories. They found that the play of the children was less constructive during the frustration period. They interpreted the results in terms of the Lewinian primitivization. CW interpreted the experiment in terms of an interference hypothesis. In an attempt to differentiate the two interpretations, five content categories were correlated with each other, using data in BDL's original paper. From the interpretations of BDL and CW, different predictions were deduced about the correlations between categories. The correlations confirm the prediction from the CW interpretation and do not confirm that from the BDL interpretation. Therefore, the CW interpretation seems to be more consistent with the data.

REFERENCES

1. BARKER, R. F., DEMBO, T., & LEWIN, K. Frustration and regression: an experiment with young children. *Univ. Ia. Stud. Child Welf.*, 1941, 18, 1-314.
2. CHILD, I. L., & WATERHOUSE, I. K. Frustration and the quality of performance. 1. A critique of the Barker, Dembo, and Lewin experiment. *Psychol. Rev.*, 1952, 59, 351-362.

DOMINANCE ATTITUDES IN PARENTS AND ADULT AVOIDANCE BEHAVIOR IN YOUNG CHILDREN¹

E. PHILIP TRAPP *and* DONALD H. KAUSLER

University of Arkansas

The fruitfulness of predicting child behavior from family variables is indicated by the trend of research in this direction (1, 4). In most of these studies, however, the effort has been to relate children's behavior to parental practices rather than to parental attitudes, mainly because of the general unreliability of attitude measurements in this area. The advantage of being able to predict children's behavior from measures of parental attitudes is obvious in terms of convenience and simplicity of data gathering.

Read (12) attempted to relate parents' expressed attitudes to nursery school children's ratings on a behavior inventory and found some relationship between attitudes of parental control and ratings on the inventory. Children of parents favoring strong control were rated with more undesirable behavioral characteristics than children of parents expressing a more liberal attitude. The ratings were the pooled judgments of the nursery school teachers.

The present study is a further investigation into the effects of parental attitudes of dominance on child behavior. Shoben (15) has developed a parent attitude scale which shows promise of being a reliable and sensitive instrument, although a recent validity study (7) indicates a limitation in its

¹ Special thanks are reserved for the nursery school staff at the University of Arkansas for their kind assistance throughout the study, and for Miss Peggy Wall, who made the observations on the children.

clinical use. The three major attitudes it purports to measure fall under the subscale headings of Dominant, Possessive, and Ignoring. Thus far, the sensitivity of the scale has been demonstrated in differentiating mothers of problem children from mothers of nonproblem children. The usefulness of the scale in differentiating less extreme cases, that is, parents of children not undergoing psychiatric treatment, is tested in the present study. More specifically, the focus of this study is on the relationship between parental dominance attitudes and adult avoidance behavior in young children.

Children of parents reflecting strong dominance attitudes are likely to develop more than the average amount of negativistic feelings toward their parents. These feelings should reinforce parent avoidance responses and generalize to other adult relationships. It is also logical to propose that parents with weak dominance attitudes might increase negativistic feelings in their children. In such cases, the child, because of insufficient guidance, is likely to experience more than the ordinary amount of frustrations and project his frustration on his parents. One manifestation of this process would be in increased parent avoidance behavior which would generalize to other adult relationships.

This proposed curvilinear relationship between parents' dominance attitudes and children's adult avoidance behavior is predicted to be in evidence only with children displaying normal personality development. In the more extreme cases of problem children, the higher anxiety drives would probably generate more variable response patterns, reflected in either excessively dependent behaviors or extreme forms of withdrawal behaviors. The net result would obscure the specific effects hypothesized here.

As a second and related problem, the present study investigated the relationship between differences in expressed parental dominance attitudes and children's behavior. It is proposed that parents with large differences in dominance attitudes between themselves would create a conflict situation for the child. Therefore, the parent-child relationships within these families are likely to be less positive than parent-child relationships within families where there is more agreement on dominance attitudes. Evidence of this less positive relationship should be indicated by increased adult avoidance behavior in their children.

These two predictions regarding the relationship between parental dominance attitudes and children's adult avoidance behavior may be stated more precisely in the following two hypotheses: (a) Nursery school children of parents scoring either relatively high or low on a dominance attitude scale will avoid adult contacts in a free-activity setting to a greater degree than nursery school children of parents scoring in the intermediate range on a dominance attitude scale. (b) Nursery school children of parents reflecting large differences in scores on a dominance attitude scale will avoid contacts in a free-activity setting to a greater degree than nursery school children of parents reflecting small differences in scores on a dominance attitude scale.

METHOD

Subjects

At the outset of the study, 20 children attending the University of Arkansas Nursery School served as Ss. By the end of the study, the normal process of attrition left completed data on 16 children. Eight Ss were males, ranging in age from 3-4 to 5-0, with a mean age of 4-2 (computed in years and months). The eight female Ss also had an age range from 3-4 to 5-0, with a mean age of 4-4. The educational and socioeconomic backgrounds of the parents were fairly homogeneous, most of them of professional middle class.

Procedure

Each parent was given the USC Parent Attitude Survey (13). The Dominant subscale served as the predictor measure. The parents were informed that the survey was part of a nation-wide study to investigate sectional differences in child-training attitudes. They seemed satisfied with the explanation and cooperated fully with the project. Anonymity was assured them by the use of coded numbers in place of their signatures. The survey was administered and collected by the director of the nursery school, who had known each parent for at least six months.

Each child was observed individually at five-minute time intervals, approximately once every 10 days through a period of 15 weeks, resulting in a total of 50 minutes of observation per child. The children were observed in the mornings between 11:00 and 11:50, the free-activity period in the nursery school program. The observations followed a prearranged schedule to insure that each child was observed throughout the entire period and never more than once at the same time of the period.

The observer was equipped with two concealed stop watches. One ran continuously for the five-minute time interval and the other watch ran only at the time the child was in contact with an adult. At the end of a day's observation, the observer made notations on the nature of the adult contacts and other qualitative remarks on the children's behavior. Since the children were accustomed to adults about them (the nursery school provides training facilities for University students), they did not show any unusual curiosity toward the observer, and, in fact, gave no indication that they were aware of being under any special form of observation. To assist in achieving this desired end, the observer spent a few days in the play area before beginning the formal observations.

RESULTS

Regarding the first hypothesis investigated, the two dominance scores for each set of parents were added together and a combined single dominance score determined. These combined scores ranged from 233 to 304.

CHILD DEVELOPMENT

The parents were divided into quartiles based on their combined scores. The first and fourth quartiles represent extreme dominance attitudes; the second and third quartiles represent intermediate dominance attitudes.

Children with parents in the extreme quartiles constituted one group of Ss (Extreme Group, $N = 8$); children with parents in the intermediate quartiles constituted a second group (Intermediate Group, $N = 8$). The two groups were found to be nearly identical in average age and average length of attendance in the nursery school. To test the related hypothesis, the mean time (in minutes) spent in adult contact was determined for each group, and the difference between group means was tested for significance by the t test. Summary data for this analysis appear in Table 1.

TABLE 1

SUMMARY OF CHILDREN'S AVOIDANCE BEHAVIOR: GROUPS DETERMINED BY COMBINED PARENTAL DOMINANCE SCORES

Group	Mean (min.)	SD	t
I Children of parents with extreme attitudes ($N = 8$)	5.7	4.6	2.3*
II Children of parents with intermediate attitudes ($N = 8$)	12.5	6.3	

* $p < .05$ (two-tailed test).

The results support the hypothesis ($t = 2.3$, $p < .05$, two-tailed test). Children with parents scoring either high or low on a dominance attitude scale displayed more adult avoidance behavior than children with parents in the intermediate range on the scale. A clearer pattern of this curvilinear relationship may be seen in a further breakdown of the data in Table 1 in terms of the quartile means. The means of the quartiles from top to bottom were 6.5, 8.6, 16.4, and 4.9, respectively.

Regarding the second hypothesis investigated, a difference score on dominance attitude for each set of parents was obtained in the following manner. For each of the 34 items on the USC Parent Attitude Survey measuring dominance, a score was recorded indicating the discrepancy between the two parents on that item. The 34 discrepancy scores were added to produce the total score for each set of parents. The parents were again divided into quartiles, this time on the basis of their difference score. The difference scores ranged from 18 to 33.

Children with parents in the first quartile (large difference between parents on dominance attitude) constituted one group of Ss; children with parents in the fourth quartile (small difference between parents on dominance attitude) constituted the second group of Ss. These two groups were likewise nearly identical in average age and average length of attendance in the nursery school. To test the related hypothesis, the mean time

(in minutes) spent in adult contact was determined for each group, and the difference between group means was tested for significance by the t test. Summary data for this analysis appear in Table 2.

The results support the second hypothesis ($t = 3.2$, $p < .05$, two-tailed test). Children with parents reflecting large differences on a dominance attitude scale displayed more adult avoidance behavior than children with parents reflecting small differences on a dominance attitude scale.

TABLE 2

SUMMARY OF CHILDREN'S AVOIDANCE BEHAVIOR: GROUPS DETERMINED BY DIFFERENCES BETWEEN PARENTAL DOMINANCE SCORES

Group	Mean (min.)	SD	t
I Children of parents with large differences ($N=4$)	5.8	1.1	3.2*
II Children of parents with small differences ($N=4$)	15.0	4.9	

* $p < .05$ (two-tailed test).

DISCUSSION

Two pertinent questions may be raised concerning the first hypothesis. The first question regards the sex of the parents. Are the dominance attitudes of mothers better predictors of their children's avoidance behavior than the dominance attitudes of fathers? The second question regards the sex pairing of parent and child. For example, are the dominance attitudes of mothers better predictors of avoidance behavior for daughters than for sons? These questions are discussed below.

An analysis of the data revealed no difference in predictive value between mothers' dominance attitudes and fathers' dominance attitudes. In both cases, the difference in avoidance scores between children with parents in the first and fourth quartiles and children with parents in the second and third quartiles was in the predicted direction and significant at the .10 level (two-tailed test). As indicated in the results section, the difference is significant at the .05 level when the dominance attitude scores are combined for parents. A further analysis also failed to disclose a closer parent attitude-child behavior relationship when parent and child are paired by sex. The dominance attitudes of both mother and father apparently affected the child about the same irrespective of the sex of the child.

A similar breakdown was made on the second hypothesis. Again the predictive value of dominance attitude scores was not enhanced either when the sex of the parents was considered separately or when pairings were made between parent and like-sex child.

The data were inspected to determine if children of parents reflecting highest difference scores on the attitude scale (Group I of Table 2) tended

CHILD DEVELOPMENT

also to be children of parents scoring at the extremes on the scale (Group I of Table 1). The inspection revealed that only one of the four children of parents in the high difference group also had parents in the extremes group. This relative lack of overlap suggests that our two hypotheses have been fairly independently confirmed by our data.

The present study supports the usefulness of the USC Parent Attitude Survey in differentiating between parents of normal children. In spite of the relatively small sample size, and the relatively homogeneous character of the sample, the scale was sufficiently sensitive to discriminate between groups at the .05 level of confidence on the two-tailed test.

SUMMARY

Two hypotheses were tested regarding the relationship between parental attitudes and children's behavior. The first hypothesis stated that nursery school children of parents scoring either high or low on dominance attitude will avoid adult contacts in a free-activity setting to a greater degree than nursery school children of parents scoring in the intermediate range on dominance attitude. The second hypothesis stated that nursery school children of parents reflecting large differences in dominance attitudes will avoid adult contacts in a free-activity setting to a greater degree than nursery school children of parents reflecting small differences in dominance attitudes.

The sample consisted of 16 children, of equal sex proportion, with an average age of four years, three months, and their parents. The USC Parent Attitude Survey was administered to the parents, and the Dominance subscale served as the predictor variable. Five-minute time samples, totalling 50 minutes of observation in all, were made on each child. The time spent by a child in adult contact during the observation period served as the dependent variable.

The results supported the curvilinear relationship predicted in the first hypothesis at the .05 level (two-tailed test). The results also supported the second hypothesis at the .05 level. Further breakdowns of the data in terms of the sex of the parent or pairing the child to the like-sex parent did not increase the predictive power of the dominance scale. The combined attitudes for both parents were more discriminating than the attitudes for either parent considered separately.

REFERENCES

1. ANDERSON, H., HANLEY, C., & HURLEY, J. Personality development in infancy and the preschool years. *Rev. educ. Res.*, 1955, 25, 453-468.
2. AUSUBEL, D. P., BALTHAZAR, E. E., *et al.* Perceived parent attitudes as determiners of children's ego structure. *Child Developm.*, 1954, 25, 173-184.

3. AYER, M. E., & BERNREUTER, R. G. A study of the relationship between discipline and personality traits in little children. *J. genet. Psychol.*, 1937, 50, 165-170.
4. BALDWIN, A. L. Child psychology. In P. R. Farnsworth (Ed.), *Annual review of psychology*. Stanford, Calif.: Annual Reviews, 1956. Pp. 259-283.
5. BORNSTON, F., & COLEMAN, J. The relationship between certain parents' attitudes toward child rearing and the direction of aggression of their young adult offspring. *J. clin Psychol.*, 1956, 12, 41-44.
6. CHAMPNEY, H. The variables of parent behavior. *J. abnorm. soc. Psychol.*, 1941, 36, 525-542.
7. GORDON, J. The validity of Shoben's Parent Attitude Survey. *J. clin. Psychol.*, 1957, 13, 154-156.
8. KOCH, H. L., The relation of certain family constellation characteristics and attitudes of children toward adults. *Child Developm.*, 1955, 26, 13-40.
9. LORR, M., & JENKINS, R. Three factors in parent behavior. *J. consult. Psychol.*, 1953, 17, 306-308.
10. MUMMERY, D. V. Family backgrounds of assertive and non-assertive children. *Child Developm.*, 1954, 25, 63-80.
11. RADKE, M. J. The relation of parental authority to children's behavior and attitudes. *Univer. Minn. Child Welf. Monogr.*, 1946, No. 22.
12. READ, K. H. Parents' expressed attitudes and children's behavior. *J. consult. Psychol.*, 1945, 9, 95-100.
13. SEARS, R., *et al.* Some child-rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.*, 1953, 47, 135-236.
14. SEWELL, W. H., MUSSEN, P. H., & HARRIS, C. W. Relationship among child training practices. *Amer. sociol. Rev.*, 1955, 20, 137-148.
15. SHOBN, E. J. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, 39, 101-148.

A STUDY OF THE VALIDITY OF PARENT ATTITUDE MEASUREMENT

DONALD A. LETON

University of California at Los Angeles

The influence of parents' attitudes on the social and emotional development of children has long been recognized. Extensive programs of parent education are being conducted through the schools, clinics, and hospitals; yet seldom, if ever, are these programs evaluated in terms of the changes produced in parents' attitudes or in children's adjustment. The present-day confusion among parents on such important variables as permissiveness, control, and discipline seems to reflect the lack of definitive research in this area.

The most extensive work in this area has been in the development of the Parent Behavior Rating Scale at the Fels Research Institute by Champney (3). This was subjected to further analysis by Baldwin, Kalhorn, and Breese (1) who have published a manual for its use. Shoben (6) and Wheeler (7) have constructed inventories for the measurement of parent attitudes and have carried out subsequent standardization and validation studies. Radke's (5) study on parental authority and Bornston's research (2) on children's aggression represent two attempts to study certain aspects of parent-child relationships. A recent study by Gordon (4) tested the validity of Shoben's scale with mothers of deaf children. Although this is not a comprehensive review of the literature on parent attitudes, it is representative of the research which has been done.

PROBLEM

Most personality and attitude scales are validated through the use of clinical or criterion groups. The items subsequently included in the scale are chosen for their power to discriminate between the criterion groups. This method does not establish item or test validity for "normal" test populations, i.e., for those subjects who would not ordinarily be classified into the criterion groups.

CHILD DEVELOPMENT

The present study represents an investigation into the validity of parent attitude measurement. The two instruments which were used in the experiment were Shoben's Parent Attitude Survey (PAS) and the writer's revision of the Minnesota Teacher Attitude Inventory (MTAI) for parents.¹ Both of these inventories were originally validated through the use of criterion groups. The attitudes of parents of problem children and of parents of nonproblem children were contrasted in the validation of the PAS. The validation studies of the MTAI contrasted teachers with desirable and undesirable teacher-pupil relationships.

The revision of the MTAI for the measurement of parent attitudes (PAI) in this study was accomplished as follows: Of the 150 items in the scale, it was found that 118 could readily be indentified as "attitudes toward children." These items were reworded by changing "pupils" to "children," "teacher" to "parent," etc. The items were then keyed and scored in the same manner as for teachers.

PROCEDURE

The PAS and the PAI were administered to a representative sample of parents in a suburb of Los Angeles, California,² obtained as follows: Alphabetical listings of the family names were obtained for all of the classrooms from kindergarten through the eighth grade at the two cooperating schools. Every fifth name was then selected from these lists to provide a sample of five families from each classroom, or approximately 20 at each grade level. The selection of names from the alphabetical listings varied with each grade level, e.g., for the first grade, 1st, 6th, 11th . . . , and for the second grade, 2nd, 7th, 12th, . . . name from each list. This procedure did not provide a sample of the total population of parents since no attempt was made to include parents whose children had not yet attained school age, nor those whose children had already completed junior high school.

In each family both parents were requested to respond to the inventories. The inventory booklets, answer sheets, and questionnaires were distributed to the homes by room mothers of the P.T.A. Parents were requested to respond to the inventories independently and asked not to discuss or compare responses until they had returned the materials. There were 283 parents who participated in the survey, of which 248, or 88 per cent, were complete pairs (*see* Table 1). The percentage of returns was approximately 78 per cent.

The mean age of the mothers in the sample was 39.2 (SD, 9.2), that of the fathers, 39.1 (SD, 6.9). The average educational attainment of these

¹ Appreciation is expressed for permission from the authors to use the Parent Attitude Survey and the Minnesota Teacher Attitude Inventory in this study.

² The writer is indebted to the Parent-Teacher Associations of Richmond Elementary School and the Junior High School of El Segundo, California, for their approval and assistance in conducting this study.

TABLE 1
CLASSIFICATION OF RESPONDENTS

<i>Category</i>	<i>Number</i>	<i>Per Cent</i>
Mothers responded, fathers nonrespondent	17	6
Mothers responded, no father present in home	6	2
Fathers responded, mothers nonrespondent or test data invalid	4	1
Mother-father pairs, test data complete (124 pairs)	248	88
Mother-father pairs, test data incomplete	8	3
Total	283	100

parents exceeded the thirteenth year (13.2) and ranged from the seventh grade level to the Ph.D. The average number of children was 2.9, with a range from 1 to 5.

RESULTS

It was first determined whether a significant relationship existed between the distributions of attitude scores for the two instruments. The correlation coefficient between the PAS and the PAI scores was found to be $-.06$. The negative coefficient is appropriate since poor attitudes are indicated by high scores on the PAS and by low scores on the PAI, and favorable attitudes, conversely. The magnitude of the relationship, however, is not significant and it is concluded that the two instruments are essentially unrelated in their measurement of parent attitudes.

The correlation between mother and father scores for the 124 pairs on the PAS was $.43$, on the PAI, $.42$, both coefficients significant at the $.01$ level.

It was next determined whether mothers had more favorable attitudes toward children than did fathers. The summary statistics are shown in Table 2. The differences between mothers' and fathers' scores on the two

TABLE 2
MEANS AND VARIANCES FOR PAS AND PAI DISTRIBUTIONS
(124 Mother-Father Pairs)

<i>Instrument</i>	<i>M E A N</i>		<i>Mean Difference</i>	<i>V A R I A N C E</i>	
	<i>Mothers</i>	<i>Fathers</i>		<i>Mothers</i>	<i>Fathers</i>
PAS	329.7	332.5	-2.8	491	366
PAI	78.5*	73.9*	-4.6	585	685

* In order to eliminate negative scores, a constant of 75 was added to each PAI score.

CHILD DEVELOPMENT

inventories were tested by a simple analysis of variance. The assumption of homogeneity of variance was found to be tenable. Although the mean differences were in favor of mothers on both instruments, the magnitude of the differences was not significant.

A total of 404 children were represented in the study, of whom 311 were of school age. It was possible to obtain adjustment ratings for 239 of the school-age children from their teachers. Ratings were not obtained for children in high school, or in private or parochial schools. The distribution of ratings is given in Table 3.

TABLE 3
DISTRIBUTION OF ADJUSTMENT RATINGS

Rating	Personal, Social, and Emotional Adjustment	N
1	Excellent	36
2	Very good	71
3	Average to above average	61
4	Average to below average	36
5	Barely adequate	18
6	Poor or unsatisfactory	17
Total		239

The mean adjustment rating for the 239 children was 2.91. There was no check on the accuracy or reliability of the teachers' ratings, and for this reason only the extreme ratings were used in further analyses. For the 17 children who received "6" ratings there were 14 families represented; for the 36 children who received "1" ratings there were 25 families. There were no families in which two siblings were at opposite extremes of adjustment; the intermediate ratings for siblings were not considered. The comparison of the two parents groups is presented in Table 4. The differences

TABLE 4
ATTITUDES OF PARENTS OF CHILDREN WITH EXTREME ADJUSTMENT

Child Adjustment Rating	N	Parent	N	P A S		P A I	
				Mean	SD	Mean	SD
1	36	Mothers . .	25	328.3	41.0	91.3	24.8
6	17	Mothers . .	14	324.9	19.9	76.9	20.7
1	36	Fathers . .	25	327.2	21.8	79.8	23.2
6	17	Fathers . .	14	326.2	26.4	76.4	28.2

between the mean scores of parents of children with good adjustment ratings and of parents of children with poor adjustment ratings were not significant for either of the attitude inventories, as tested by analysis of variance.

The data permitted a test of a widely accepted concept in child development that parental agreement has a favorable effect on a child's adjustment and that parental disagreement has an unfavorable effect. The differences between the individual mother and father scores on the PAS and PAI were computed for the 39 pairs at the two extremes of child adjustment. The differences on both instruments were then combined into a single discrepancy score to determine whether the parents of the poorly adjusted children had a wider discrepancy, or "difference-in-attitude," score than did the parents whose children received the best adjustment ratings.

It can be seen in Table 5 that there is a greater discrepancy between the attitude scores of parents of poorly adjusted children than between parents of well adjusted children. The difference is significant ($t = 6.23$, $p < .001$).

TABLE 5
ATTITUDE DISCREPANCY SCORES OF MOTHER-FATHER PAIRS

<i>Child Adjustment Rating</i>	<i>Number of Pairs</i>	DISCREPANCY SCORES	
		<i>Mean</i>	<i>SD</i>
1	25	36.5	23.8
6	14	55.4	23.9

SUMMARY

Shoben's Parent Attitude Survey and a revision of the Minnesota Teacher Attitude Inventory were administered to a representative sample of parents of children in kindergarten through the eighth grade. A summary of the findings follows:

1. There was no significant relationship between scores on the two attitude inventories.
2. There was a significant similarity between mother and father attitudes toward children within a given family.
3. Although mothers obtained more favorable scores than did fathers on both instruments, the differences were not significant.
4. There was no significant difference between the attitude scores of parents whose children received ratings of "excellent adjustment" and those whose children received "poor adjustment" ratings.

CHILD DEVELOPMENT

5. There was a wider disagreement in attitudes between mothers and fathers of poorly adjusted children than between mothers and fathers of well adjusted children.

REFERENCES

1. BALDWIN, A. L., KALHORN, J., & BREESE, F. H. Patterns of parent behavior. *Psychol. Monogr.*, 1949, 63, No. 4.
2. BORNSTON, F. L., & COLEMAN, J. C. Relationship between certain parents' attitudes toward child rearing and the direction of aggression of their young adult offspring. *J. clin. Psychol.*, 1956, 12, 41-44.
3. CHAMPNEY, H. The measurement of parent behavior. *Child Development*, 1941, 12, 131-166.
4. GORDON, J. E. The validity of Shoben's Parent Attitude Survey. *J. clin. Psychol.*, 1957, 13, 154-156.
5. RADKE, MARIAN J. The relation of parental authority to children's behavior. *Univer. Minn. Inst. Child Welf. Monogr.*, 1946, No. 22.
6. SHOBN, E. J. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, 39, 101-148.
7. WHEELER, M. E. Preliminary standardization of a parent-child relationship scale. *Dissertation Abstracts*, 1955, 15.

DEVELOPMENTAL STATUS OF TWO GROUPS OF INFANTS RELEASED FOR ADOPTION¹

D. BRUCE GARDNER and MARYBELLE K. SWIGER
Child Development Research Laboratories, Iowa State College

A research project undertaken by the Child Development Research Laboratories in 1955 had as its central objective the investigation of the effects of home management house experience on the development of infants. A discussion of the theoretical bases of the research has been presented earlier (3), and will not be included here.

In the over-all research project, various measurements of development and adjustment are being made at each of seven specified age levels during the first two years of life. Infants who live in a college home management house are compared, with respect to these measurements, with infants who live in private foster homes, and also with infants who live in their own parental homes from the time of birth. In order to make valid and meaningful comparisons of these three groups of infants, and relate these comparisons to the experimental conditions which define each of the groups, it is of fundamental importance to establish the degree of comparability of the groups before they are exposed to the unusual mothering experiences. This report describes the developmental status of the subjects at the time they were first inducted into the project, prior to being subjected to the experimental conditions under investigation.

There are obvious practical justifications for making the comparisons reported here, for the reasons explained above. In addition, there are possibly some theoretical implications which go beyond the demands of our project. The fundamental question of differences in genetic endowment

¹ Journal Paper No. J-3447 of the Iowa Agricultural and Home Economics Experiment Station, Ames, Iowa, Project No. 1299. Acknowledgement is due the Elizabeth McCormick Memorial Fund for partial financial support, and the staff of the Iowa Children's Home Society for their cooperation.

CHILD DEVELOPMENT

between infants whose mothers release them for adoption and infants who are born into complete families is raised here. The existence of constitutional differences as a function of legitimacy of birth, and the degree to which any such possible differences may be modifiable by early mothering experiences, will remain open questions in the absence of definitive research on the specific variable of legitimacy. However, the manner of selecting subjects for the groups in this study, and the over-all experimental design involved, may enable us to make some tentative statements on the role of genetic factors in the development of adopted children.

DESCRIPTION OF THE SUBJECTS

In this report Group A includes 37 infants who subsequently are subjected to residence in a home management house for a period of one college quarter, or about three months, prior to adoption. Group B includes 44 infants of comparable background to those in Group A, but who reside in private homes with foster parents for a period of time before being adopted. Infants in both of these groups are born to unmarried mothers and released for legal adoption purposes. The selection of infants for the home management house was made by the staff of the Iowa Children's Home Society, which agency supervised the placement and adoption of the infants. The only basis of discrimination between Groups A and B was the practical consideration of availability of infants at the time each college quarter began.

A third group, C, is made up of 47 infants who were born into complete families and who will, presumably, be subjected only to "normal" mothering experiences in their own natural homes.

All of the subjects were born in the State of Iowa, and with few exceptions they were born in Des Moines hospitals. Infants with any obvious congenital defects such as spastic paralysis, malformations of limbs or body organs, micro- or hydrocephaly, or serious sensory defects, were excluded. Only full-term infants were accepted, the criterion for "full-term" being a minimum birth weight of 5½ pounds.

TABLE I
CHRONOLOGICAL AGE IN DAYS AT TIME OF EXAMINATION

	M A L E S			F E M A L E S			T O T A L		
	N	Range	Mean	N	Range	Mean	N	Range	Mean
Group A	21	7-59	27.1	16	4-72	29.0	37	4-72	27.9
Group B	21	8-64	30.6	23	7-75	27.3	44	7-75	28.9
Group C	25	21-92	39.5	22	18-57	34.2	47	18-92	37.0
Total	67	7-92	32.9	61	4-75	30.2	128	4-92	31.6

The sex distribution within groups was controlled as rigidly as practicable within the limits of availability of infants, and the desirability of including as many subjects as possible. However, the rigid control of this variable was thought to be a less vital consideration than maintenance of reasonably good sized groups, since sex differences were not the prime variable under consideration. Table 1 reveals the degree to which the three groups and the total were comparable with respect to relative proportions of males and females.

In attempting to determine the comparability of the three groups, a most important consideration is that of the age at which the measurements were taken. Efforts were made to control the examination age, by limiting the enrollment of subjects to those under three months of age, and by attempting to test them as near as possible to the 30th day following birth. For reasons of scheduling and practical considerations in the placement of infants in the home management houses and foster homes, it was impossible to control this variable as well as would have been desired. Table 1 provides data showing the mean ages and age ranges at time of examination, for the three groups and also for sexes.

The mean ages for males and females were not significantly different, either within any group or for the combined groups. However, a comparison of the three groups without regard to sex revealed that Group C infants were significantly older than were those in the other two groups ($F = 3.686$, significant at the .05 level with 2 and 127 *df*). This difference necessitated a statistical control of the age variable in order to analyze the data for comparative developmental status of the three groups.

Table 2 provides information on two measurements of gross body development at birth, those of weight and length. The sex difference in birth length approaches significance at the .05 level ($F = 3.788$ with 1 and 123 *df*). Analysis of variance was used to test the significance of differences among groups, as well as between sexes, for both weight and length at birth. However, none of the F ratios, other than that for sex difference in length, was near significance at the .05 level, and they are not shown.

TABLE 2
PHYSICAL MEASUREMENTS AT BIRTH

	M A L E S			F E M A L E S			T O T A L		
	N	Weight (kg.)	Length (cm.)	N	Weight (kg.)	Length (cm.)	N	Weight (kg.)	Length (cm.)
Group A	21	3.40	50.84	15	3.31	50.67	36	3.36	50.77
Group B	20	3.41	51.98	22	3.44	51.35	42	3.42	51.65
Group C	24	3.51	51.57	22	3.20	49.68	46	3.37	50.66
Total	65	3.44	51.46	59	3.32	50.55	124	3.39	51.03

Similarly, interaction effects of sexes and groups were examined with the *F* test, and were found to be negligible for both variables.

A thorough medical examination by the consulting pediatrician was made prior to induction of each infant into the research. Since this screening examination eliminated any infants with marked pathological conditions, the net effect was to restrict the range of variation in physical status among the subjects. The nature of the medical examination reports, plus the fact that the average age of infants in Groups A and B (10 days) was considerably lower than that for Group C (31 days) when the medical records were taken, made statistical comparison of the groups difficult and hazardous.

A simple tabulation of positive medical findings in each of the areas included in the pediatric report (including body temperature, skin, sensory organs, nose and throat, abdomen, heart, lungs, genitals, body posture, and general nutritional status) revealed the following incidence of anomalies: Group A, 46; Group B, 40; Group C, 33. In terms of percentages, Group A, which included 30 per cent of the cases, accounted for 39 per cent of the anomalies. Group B, with 35 per cent of the cases, contributed 33 per cent of the anomalies. Group C, with 35 per cent of the cases, contributed 28 per cent of the anomalies.

Unfortunately, statistical tests of differences for these proportions would be relatively meaningless, for at least two reasons. First, the wide range of items listed above and lumped together under the heading of "anomalies" cut across many dimensions of observation and measurement, and they are certainly not of equal significance medically. Second, such variables as general nutritional status, which was more frequently noted as a positive finding in Groups A and B than in Group C, are probably a function of age. In general, there is nothing in the medical data to indicate that one of the groups is significantly inferior or superior to the others.

It might be expected that differences in developmental status among the three groups could arise from hereditary variation, if it could be assumed that there are some heritable characteristics associated with out-of-wedlock birth. However, there seems to be little justification for making any such assumption in the absence of any evidence that the tendency to out-of-wedlock parentage is associated with genetic structure. Lacking such evidence, we are assuming that genetic effects on developmental status are distributed no differently in one group than in the others. The data reported here are, in one sense, a test of the validity of this assumption, with respect to a limited number of measurements of developmental status taken within a relatively narrow range of chronological age.²

² Admittedly this assumption may not be fully warranted since information on parentage in Groups A and B is meager. Furthermore, the age of mothers at time of subjects' birth is lower, on the average, for Groups A and B than for Group C. The extent to which this factor may have influenced the constitutional make-up of the subjects cannot easily be determined.

MEASURING INSTRUMENTS AND EXAMINATION PROCEDURE

The variables of physical and psychological status are represented by a variety of specific measurements. Anthropometric measurements include body weight, length, head circumference, chest circumference, calf circumference, bi-iliac (hip) width, and subcutaneous leg tissue. Standard measuring instruments were used to obtain these data. Steel tapes, a standard measuring board, and basket-type scales were used to obtain circumferences, length, and weight respectively. Bi-iliac width was measured with large sliding calipers; small spreading calipers were used to measure subcutaneous tissue in the left leg. Calf circumference was also measured on the left leg.

Body measurements were taken by a nurse and an assistant, both of whom received preparatory instruction in techniques of anthropometric measurements with infants. Each measurement was made twice, the second time by the assistant, and the average of the two readings was computed. If at any time a sizable discrepancy was noted between the two readings, a third one was taken and averaged with the nearer of the first two.

The Gesell Developmental Schedules were administered by a psychologist, following the procedure outlined by Gesell and Amatruda (2), and using the standard examination materials. The scales were designed to estimate developmental status in each of four areas (language, motor, adaptive, and personal-social), as well as providing a total score. Since the number of items in each area is small, and the precise meaning of the areas somewhat open to question, only a general developmental level was computed for purposes of this report.³

The California Infant Scale of Motor Development was also administered by a psychologist. This scale was developed by Bayley (1) for measuring motor skills and coordination. Each item is age-scaled, and the items range from .2 months to 54.5 months. A point score is computed, based on the actual number of items passed. This is then converted to a sigma score, on the basis of mean and standard deviation figures provided for each age level by Bayley.

Each subject was examined individually, and 30 to 40 minutes were required to complete the psychological and anthropometric measurements. Subjective notes of the examination behavior of the infant were recorded by each of the two psychologists present at the examination. The sequence

³ A somewhat more precise quantification of scores on the Developmental Schedules than that suggested by Gesell and Amatruda (2) was desired for purposes of statistical analysis. Therefore, the scales were treated as if they were a psychometric test, and the child's performance level was computed as follows: key age was determined by noting the highest test level at which over half the items were passed. This estimate of developmental age was corrected by subtracting a number of days proportionate to the number of failures at that level and at lower levels. If items were passed at higher levels, a number of days of developmental age proportionate to the number of passes was added to obtain developmental age level. This figure was divided by chronological age and multiplied by 100 to obtain the general developmental quotient.

CHILD DEVELOPMENT

of tests was always the same: the Developmental Schedules were followed by the Scale of Motor Development, after which the anthropometric measurements were taken. An examination room at the Iowa Children's Home Society was used for Groups A and B, and an examining room at the pediatric clinic was used, by the same team of examiners, for Group C. Each infant was accompanied in the testing room by his mother, or foster mother. The length of time required to complete the induction of all subjects into the project was approximately one year.

RESULTS

The basic findings, in the form of mean scores in each of the variables, are presented in Table 3. Tests of significance of differences between sexes revealed that the mean scores favored males in body length ($F = 5.205$, significant at the .05 level), and in head circumference ($F = 5.844$, significant at the .05 level). None of the other six differences was significant at that level; however, the difference in weight, which also favored males, was significant at the .10 level ($F = 3.053$). The remaining anthropometric measurements, with the exception of leg tissue, favored males but the differences were small and not statistically reliable.

TABLE 3
MEAN SCORES AND MEASUREMENTS FOR GROUPS AND SEXES

Variable	Group A Mean N	Group B Mean N	Group C Mean N	Males Mean N	Females Mean N	Total Mean N
Weight (kg.)	3.94 37	4.01 44	4.48 46	4.29 66	4.02 61	4.16 127
Length (cm.)	53.44 37	54.09 44	54.77 46	54.79 66	53.45 61	54.15 127
Head (cm.)	36.66 35	36.46 43	36.94 44	37.07 62	36.30 60	36.69 122
Chest (cm.)	34.09 36	34.26 43	35.19 45	34.86 64	34.21 60	34.55 124
Calf (cm.)	11.83 35	12.04 43	12.78 44	12.27 62	12.22 60	12.25 122
Hip (cm.)	8.59 37	8.68 43	8.84 46	8.76 66	8.67 60	8.72 126
Leg tis. (mm.)	15.25 37	15.71 39	17.38 44	16.03 62	16.34 58	16.18 120
Motor (sigma)	1.48 35	1.41 44	0.58 46	1.06 66	1.20 59	1.13 125
DQ	135 37	136 44	93 47	114 67	126 61	120 128

A statistical comparison of the three groups with respect to physical measurements necessarily had to take into account the discrepancy in chronological age, since it might be assumed that these variables are correlated with age. Therefore, analysis of covariance was used to test for significance of differences among group means adjusted for age. Table 4 presents the basic data derived through the covariance analyses.

The consistently significant correlation of each of the variables with age is one indication of the justification for resorting to the covariance analyses. Comparisons of the adjusted mean scores for groups, shown in Table 4, with the unadjusted mean scores reported in Table 3 reveals the tendency for the age adjustment to bring the mean scores for groups much closer to each other. The F test here represents the ratio of the adjusted mean square for variance between groups to the adjusted mean square for variance within groups. Only in the case of calf circumference is the F ratio significant ($F = 3.113$, significant at the .05 level with 2 and 120 df).

TABLE 4
MEAN SCORES AND MEASUREMENTS ADJUSTED FOR AGE

<i>Variable</i>	<i>r*</i>	<i>Group A</i>	<i>Group B</i>	<i>Group C</i>	<i>F</i>
Weight (kg.)	.68	4.06	4.10	4.29	1.494
Length (cm.)	.66	53.90	54.43	54.14	.430
Head (cm.)	.69	36.86	36.66	36.67	.288
Chest (cm.)	.77	34.39	34.47	34.69	.483
Calf (cm.)	.66	11.96	12.14	12.53	3.113
Hip (cm.)	.65	8.72	8.74	8.67	.098
Lcg tis. (mm.)	.57	15.59	16.18	16.76	1.530
Motor (sigma)	-.27	1.17	1.32	0.64	1.154
DQ	-.64	126	129	107	2.070

* Correlation of each variable with chronological age.

Statistical summaries of the scores obtained with the California Infant Scale of Motor Development and the Gesell Developmental Schedules are also presented in Table 4. It should be noted that in both of these sets of data scores are reported in the form of derived indices, sigma score and developmental quotient respectively. Since the derivation of these scores from raw scores involves computing a ratio of performance to chronological age, the use of the derived index should have the effect of controlling the age variable. However, several characteristics of the data suggest the presence of a negative correlation between derived scores and chronological age. After computing the correlation coefficients and noting the extent of this relationship, it seemed appropriate again to make use of the covariance technique to control the age variable.

Adjusted mean scores for motor development and developmental quotient are shown in Table 4. The F ratios indicate that the differences are not significant for either variable.

DISCUSSION

The one measurement in which significant differences among groups remain after the age adjustment is made, that of calf circumference, is not easily interpreted in relation to the rest of the data. The absolute difference between any two means is not great: the largest, between Groups A and C, is slightly more than half of one centimeter. This could not be said to be of fundamental significance as a basis for distinguishing between two groups. Presumably, in any large series of statistical tests of differences among given groups, the chances are fairly good that a "significant" *F* ratio may occur in the absence of any very important real difference.

Nevertheless, the primary purpose of the analyses reported here was to establish a base line of comparison of the three groups prior to experimental treatment in the larger longitudinal study. Since this difference in calf circumference, however small, does appear to be associated with the selection of subjects for the groups, it should be of interest to note, in subsequent analyses of data obtained at older ages, whether this difference remains over a period of time allowing for maturation and more extensive use of the body in walking and other motor activity.

So far as the remaining variables are concerned, it may safely be assumed that the three groups are not fundamentally different with respect to the variables we have measured. The logic implicit in the statistical tests employed here does not allow us to say that the subjects in the different groups are all members, drawn at random, of some larger population. Speaking conservatively, however, it is evident that no one group differs consistently or systematically from the other two.

There are both practical and theoretical implications of these findings. From the standpoint of the larger research project, this serves to increase our confidence in the validity of comparisons to be made following experimental treatment of Groups A and B. The question may be raised, of course, about underlying genetic differences which might not have had time to appear through normal maturation processes during the first few weeks after birth. The best answer to this question, so far as the longitudinal design is concerned, will probably become available through a careful use of the foster home group data. Since these subjects are selected in the same manner and from the same sources as are the home management house infants, any underlying differences in genetic endowment between the home management house group (A) and the control group (C) should also appear between the foster home group (B) and the control group (C).

One interesting aspect of the results certainly is the striking negative correlation between developmental quotient and chronological age. A possible explanation for this finding lies in the nature of the Gesell Schedules and the normative data available on these scales for the early weeks of infancy. For the four-week level, the earliest age at which norms are available,

the items do not appear to discriminate well among infants whose chronological ages range from 0 to 30 days. That is, the infant who is one week old frequently is observed to "pass" as many of the four-week items as the infant whose age is three or four weeks. This failure to discriminate finely results in a distortion which is mathematically amplified in the process of computing a DQ score with such young subjects. The tendency of the youngest subjects to receive disproportionately high DQ scores undoubtedly contributes to the strong negative relationship between CA and DQ. It would seem appropriate to suggest that research workers using the Gesell Schedules with very young infants might, if it is important to compute DQ scores, add a constant both to the numerator and the denominator in the DQ formula. The addition of, say, 60 days should tend to reduce this apparent distortion considerably.

SUMMARY

The longitudinal investigation of the effects of noncontinuous mothering on human development necessitated comparison of the subjects prior to exposure to the different kinds of mothering under investigation. This report describes the developmental status of three groups of infants at the time they were first inducted into the research, prior to any experimental "treatment."

Two groups of infants, A and B, were born to unmarried mothers and released by them for adoption. A third group, C, was composed of infants born into complete families under normal circumstances. The proportion of subjects of each sex in each group was kept relatively even. Efforts were made to examine all infants at or near one month of age; however, Group C infants were slightly older, on the average, than those in the other two groups. This difference made it necessary to apply statistical control to the age variable in order to compare the groups.

For each group, data were obtained on a series of anthropometric measurements, motor development as rated with the California Infant Scale of Motor Development, and general developmental status as rated with the Gesell Developmental Schedules.

Significant sex differences, favoring male subjects, were noted in body length and head circumference. Analysis of the data by groups revealed that Groups A, B, and C were roughly comparable in developmental status at the time of induction into the research. In only one measurement did a significant difference appear, that of calf circumference. It would seem as plausible to interpret this as a chance variation, which could occur in such a series of tests of significance, as to assume that it is a reliable index of fundamental differences among the groups, since there is little else in the data to support this specific finding.

Significant negative correlations of developmental quotients with chronological age were noted, and were attributed largely to the inadequate dis-

CHILD DEVELOPMENT

criminatory powers of the scales during the early weeks, as well as the very small CA factor used in the denominator of the DQ formula. A mathematical correction for the distortion resulting from these two conditions was suggested.

REFERENCES

1. BAYLEY, NANCY. The development of motor ability during the first three years. *Monogr. Soc. Res. Child Develpm.*, 1936, 1, No. 1.
2. GESELL, A., & AMATRUDA, CATHERINE. *Developmental diagnosis*. (2nd Ed.) New York: Hoeber, 1954.
3. PEASE, DAMARIS, & GARDNER, D. B. Research on the effects of non-continuous mothering. *Child Develpm.*, 1958, 29, 141-148.

VARIABLES IN EARLY DISCRIMINATION LEARNING:
II. MODE OF RESPONSE AND STIMULUS
DIFFERENCE IN THE DISCRIMINATION
OF TONAL FREQUENCIES¹

WENDELL E. JEFFREY

University of California, Los Angeles

When one uses adult Ss in a discrimination experiment, the indicator response rarely causes any problems. One can ask that any of a number of types of judgments be made, in a number of different ways, without substantially increasing the difficulty of the task. With young children, however, the response required can make a critical difference in the speed of learning, or indeed, may make the difference between learning and not learning. For example, Jeffrey (3) demonstrated that children who could not learn to apply distinctive labels to stimuli differing in spatial orientation could nevertheless learn to press buttons toward which the stimuli (stick figures) could be said to be pointing. It seemed important to investigate further the various factors that contribute to this early discrimination learning. The spatial discrimination problem previously used was a relatively easy one. In order to get more reliable measures of the contributing factors, it was advisable to find a stimulus dimension that was more difficult and one that was also more readily manipulable.

The tonal frequency dimension (pitch) met these criteria. Preliminary research indicated that kindergarten children had difficulty in responding differentially, and that the words, "high" and "low," meant nothing to children when applied to musical notes, indicating that the stimuli were

¹ This research was supported by grant M-1498 from the National Institute of Mental Health, Public Health Service, and by grant G-4414 from the National Science Foundation. I wish to express my appreciation to Mrs. Eleanore Cornberg of the University School Kindergarten for her cooperation in this research.

not already verbally predifferentiated. Although adult humans tend to associate low tones with left, whether the button on the left was to be associated with a high or low tone made no difference to the children. Nevertheless, it seemed reasonable to believe that a response might be found that would be readily learned and possibly facilitative in the acquisition of those responses previously demonstrated to be very difficult to learn.

A response to pitch that seemed to have some of the feedback characteristics of the button pressing response to the stick figures in the research previously noted was that of singing a matching tone. Therefore, this task was chosen for one group of Ss. Because it was found that it is difficult for children to sing a particular note at will, another group was trained on an analogous task, viz., to produce matching tones on the piano as their indicator response. Button pressing, with buttons oriented to the right and left of S, provided a control task performed by a third group and served as a criterion test task after training with singing or with the piano.

Inasmuch as the discrimination of pitch proved to be a very difficult task, it was thought that the variable of stimulus difference should also be manipulated. The data of psychophysics and studies of stimulus generalization indicate that when stimulus differences are large discrimination learning is likely to be easy. There is, however, little information regarding the transfer of training from learning an easy discrimination to learning a more difficult one, when difficulty is defined in terms of separation along a stimulus continuum.

It could be said that the first responses a child has to learn in life are of the sort that define or orient the child to the continuum (1) and that gross stimulus differences help achieve this end. Once the continuum is defined, subsequent learning with regard to stimuli within the continuum would be expected to be facilitated. The present experiment first presented the children of the experimental groups with an opportunity to define the pitch continuum, then measured their success by means of a task otherwise beyond their ability, and finally made the latter task even more difficult by requiring closer discrimination. Two hypotheses were being tested: (a) that practice on the two separate pretraining tasks involving the production of tones to match the stimulus tones would transfer to the difficult task of pressing the left button after a low stimulus tone and the right button after a high stimulus tone; and (b) that training with a pair of tones separated by 1024 cycles would facilitate subsequent discrimination of a pair of tones separated by only 128 cycles.

METHOD

Subjects

The Ss were 21 children, 11 boys and 10 girls, from the University of California Senior Kindergarten. They were divided into three groups of seven each. The mean ages were 5 years, 6 months for Groups I and III

and 5 years, 7 months for Group II. The range for Group I was 5-2 to 5-10, for Group II, 5-4 to 6-0, and for Group III, 5-4 to 5-11.

Apparatus

Two tape recorders were used, one to present children's music, which was an incentive and reinforcer, and one to present the discriminative stimuli. Further apparatus programmed push buttons in such a way that if the correct button was pressed within three seconds following the occurrence of a discriminative stimulus the music continued to play. If the incorrect button was pressed, the music terminated. The music could be heard again only upon making a correct response following the presentation of the next discriminative stimulus. The response buttons were approximately 14 inches apart on a small board that could be placed on a table in front of *S*, or on the floor to be operated with *E*'s feet. The discriminative stimuli and the music were delivered through separate earphones worn by *S*. The tone was sufficiently loud to cause no confusion when it occurred along with the music. The results of the piano response group, presented below, confirm this.

Stimuli

Two series of piano tones were recorded. The first series, Series A, consisted of two tones, C_1 and G^3 (128 and 1152 cps.), recorded serially at 10-second intervals in the order HLHHLHL. A second series, Series B, consisting of the tones, C and G (256 and 384 cps.), was recorded in the same manner.

Procedure

There were three groups of *Ss*. Group I was trained only on the button pressing response; Group II was required to match the discriminative stimuli vocally prior to being tested on the button pressing response; and Group III learned to produce matching tones on the piano before being tested on the button pressing response. All of the above training and testing was done with Series A. Any *S* who met the criterion on the button response with Series A was tested for transfer to Series B. A criterion of 10 successive correct responses was used throughout.

Each *S* in Group I was brought into the experimental room and seated at a low table upon which were push buttons. Earphones were placed on his head and he was permitted to listen to music for approximately one minute. The music was then terminated and *S* was told that he would hear tones in the earphones and *E* sang representative tones and indicated which button was to be pressed for each tone in order to continue the music. The music was then turned on again and Series A (128 and 1152 cps.) was presented. Transfer of the button pressing response to Series B (256 and 384 cps.) stimuli was tested if the criterion of 10 successive correct discriminations on Series A was met within 72 trials. An earlier study indi-

CHILD DEVELOPMENT

cated that if *Ss* had not learned in 72 trials they were not likely to learn even if they could be kept in the situation for a longer period.

The *Ss* of Group II were instructed similarly to Group I except that they were asked to reproduce the Series A stimuli by singing or humming the tone that they had heard rather than pressing a specific button. Examples were given by *E* and the children were helped to move their voices to the correct tone by *E*'s singing along with them. On meeting the criterion on the singing response, these *Ss* were presented with the button pressing task with the Series A stimuli. If they met the criterion of 10 successive correct responses on the button task, transfer of this response to Series B stimuli was also tested.

Each *S* in Group III was seated at a piano in the experimental room rather than at the table. After listening to the music, he was shown that two keys of the piano were marked with red tape, and was told that his task was to match the tone he heard through the earphones by pressing the correct key. As with the other groups, this task was demonstrated by *E*'s singing representative tones. Subsequent to meeting the criterion on the Series A stimuli, performance on the button response was tested. Those *Ss* who met the criterion on this task were then presented with the Series B stimuli.

RESULTS

Of first consideration in the results is the fact that only one of seven *Ss* in Group I learned to press the buttons differentially to the Series A stimuli within 72 trials. One *S* in this group refused to continue after 48 trials. This failure to learn or to continue meant a score of zero or, more exactly, an infinite score. Thus, the data can be best summarized in a pass-fail contingency table (Table I), with passing defined as meeting a criterion of 10 successive correct responses within the 72 trials.

Considering first the button pressing criterion responses to the Series A (128 and 1152 cps.) stimuli, only one *S* met the criterion in Group I, five *Ss* met the criterion in Group II, and seven *Ss* met the criterion in Group

TABLE I
NUMBER OF CHILDREN MEETING CRITERION OF 10 SUCCESSIVE
CORRECT DISCRIMINATIONS IN THE ALLOTTED
NUMBER OF TRIALS

	N	Series A C ₁ - G ³	Series B C - G
Group I (buttons)	7	1	0
Group II (singing)	7	5	1
Group III (piano)	7	7	1

III. Group II Ss had had from a minimum of 24 to a maximum of 72 prior trials attempting a differential singing response. The median was 24 responses, four Ss seemingly being able to sing different notes reliably to the stimuli although not necessarily reproducing them exactly as presented. The general inability of the children in Group II to come even close to reproducing the stimuli made it difficult for *E* to score their responses. One *S* never produced differential tones that were in any way acceptable. The other Ss were reinforced for anything that sounded like a differential response in the hope of shaping their subsequent responses. Whether two observers, however, could have agreed on the acceptability of each response is not known. The fact remains that of the six Ss who, in *E*'s opinion, were learning to make differential singing responses, five subsequently learned to make the button pressing response.

Group III had from 10 to 24 trials making the piano responses prior to the test situation. Four of these Ss met the criterion with the piano immediately after instructions. The other three Ss took 9, 12, and 15 trials to meet the criterion with the piano. Transfer to the button response was perfect in all seven cases. Thus, for Group III, the criterion for the button pressing to Series A stimuli was reached by four Ss without experiencing failure and with only 10 pretraining trials. In the other three cases it took 19, 22, and 25 trials, 10 of these trials in each case being criterion trials for the piano response.

The differences in performance on the transfer task (button pressing) can best be evaluated by the Fisher exact probability test which tests the hypothesis as to whether two groups differ in the proportion with which individuals fall into two classifications. The difference between Groups I and III is significant with $p < .01$. The difference between I and II has an exact probability of .0512. Group II does not differ significantly from III. Thus, the hypothesis that such pretraining would be facilitative is supported. These findings are almost identical to those obtained in a preliminary study reported elsewhere (2).

The hypothesis regarding transfer from training with the Series A stimuli, involving a difference of three octaves and a fifth between stimuli, to the closer discrimination of the Series B stimuli, a difference of only a fifth, was not supported (Table 1). In Group I only one *S* learned the Series A discrimination, and he did not learn the Series B discrimination. In Group II, although five Ss learned the Series A discrimination only one *S* demonstrated any transfer to the Series B stimuli. In Group III, even though all Ss learned the button pressing response to Series A, again only one *S* showed transfer to the Series B stimuli.

DISCUSSION

Impossible as it may seem, 5½-year-old children who can sing their favorite song (e.g., Davy Crockett) with quite acceptable intonation, are

CHILD DEVELOPMENT

not able to make a differential button pressing response to pitch stimuli without certain pretraining. They also do very poorly at singing tones that match presented stimuli even when given all possible aid by *E*. The degree to which the *Ss* had difficulty with the button pressing task needs to be emphasized. The instructions were very specific and an *S*, after instructions, might make four or five correct responses in succession only to apparently suddenly become confused and no longer able to remember which button was appropriate for each tone. His performance then dropped to chance level. In greatest contrast to this performance is the performance of the piano response group which, on the whole, had very little difficulty making what in practically all aspects was an identical response except for the feedback it produced. This feedback apparently accounted then for not only faster learning on the piano response task but also the high positive transfer to the button pressing task. The fact that the piano response group learned so readily also indicates that the incentive conditions were sufficient.

There may be some question about the stability of this training, however, over longer periods of time. Therefore, before one can go very far in speculating about the processes involved, it is important to know whether the button pressing response will persist over a period of days, rather than just over a number of trials. If it is only temporary, it is highly probable that its permanence could be increased by further training. It is of theoretical importance to know, if more training is necessary to assure permanence, whether this added training would be more beneficial if done on the piano or with buttons.

In similar perceptual situations with children, psychologists have often been unable to determine whether failure to make a discrimination is due only to *S*'s inability to make an overt response or to his inability to make a discrimination of any kind. In this instance the success of the piano group suggests that *Ss* are able to make the discrimination and therefore their failure to sing must result from the lack of differentiation of the singing response. On the other hand, it is to be noted that the button pressing group could not respond differentially. The question arises as to whether this group could discriminate but could not make the response or whether the basic perceptual discrimination occurred as the result of making the piano and singing responses. Such an hypothesis is supported by the fact that the singing group subsequently performed the button pressing response even though their singing responses had been painfully inaccurate. These results apparently may be explained in several ways. First, one might assume that the feedback of the singing and piano responses is important in some way. Certainly, it is hard to think of a tone or a tone discrimination without making a covert response. Both Mandler (5) and Kanfer (4) have accentuated the importance of such responses in discrimination learning, and to require a singing response should help implement this process. One would not necessarily predict, however, that the piano response would produce the appropriate mediational responses. A second hypothesis seems

more plausible here. One might assume that the correctness of a response is accentuated by the immediate comparison between *stimulus tone* and the response-produced tone. This would provide a relatively easy same-difference discrimination and should help accentuate the continuum involved as well as provide a more direct form of reinforcement. A third explanation, appropriate at least for the piano response group, is the fact that the piano response shortens the interstimulus interval by effectively extending each stimulus in time. The relevance of this variable is, of course, subject to test and work in this direction is now in progress.

Another source of difficulty in the discrimination may have arisen from the fact that there were common overtones between the two notes within each series, inasmuch as the piano was used as a stimulus source. This variable will also be assessed in future research.

SUMMARY

In order to study some factors basic to early discrimination learning in children, the tonal frequency dimension was chosen as the source of stimuli. Both the type of indicator response and the difference between stimuli were varied.

The Ss were 5½-year-old children. The criterion task was to learn to push two buttons, oriented to the left and right, differentially to two tones separated by three octaves and a fifth. Group I had training only on this task. Prior to testing on the criterion task, Group II Ss were required to match the same two tones by singing, and Group III produced matching tones on a piano. Although Ss in Group II had difficulty making a match, they did seem to make differential responses, and when transferred to the button pressing criterion task, these Ss performed significantly better than those in Group I. The Group III Ss, who were required to produce matching tones on a piano, learned this quite readily and showed almost perfect transfer to the criterion task. Thus, Groups II and III differed significantly from Group I but not from each other.

All Ss who met the criterion on the button pressing task were subsequently required to make this same response to stimuli differing by only five-eighths of an octave. There was no evidence of transfer to these new stimuli in any of the groups. Suggestions for further research are discussed.

REFERENCES

1. EHRENFREUND, D. An experimental test of the continuity theory of discrimination with pattern vision. *J. comp. Psychol.*, 1948, 41, 408-422.
2. JEFFREY, W. E. An experimental study of some factors in early discrimination learning. *Amer. Psychologist*, 1957, 12, 392. (Abstract)

CHILD DEVELOPMENT

3. JEFFREY, W. E. Variables in early discrimination learning: I. Motor responses in the training of a left-right discrimination. *Child Develpm.*, 1958, 29, 269-275.
4. KANFER, F. H. Perception: Identification and instrumental activity. *Psychol. Rev.*, 1956, 63, 317-329.
5. MANDLER, G. Response factors in human learning. *Psychol. Rev.*, 1954, 61, 235-244.

